

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2002, 08:28:48 ; Search time 3090.94 Seconds
(without alignments)
11947.068 Million cell updates/sec

Title: US-09-914-001-1

Perfect score: 2736

Sequence: 1 atggcgccgagctcccgctc.....caggggagtgatgtgagcag 2736

Scoring table:

IDENTITY NUC
Gapol 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	483	17.7	500	9	AW065905
2	411.6	15.0	638	10	BE415676
3	407	14.9	841	10	BF625060
4	384	14.0	717	10	BM073464
5	358	13.1	558	10	BE345465
6	324	11.8	761	10	BI178113
7	309.8	11.3	590	10	BE455437
8	308	11.3	418	9	AI586713
9	296	10.8	507	10	BE475748
10	295.4	10.8	613	9	AW330561
11	292	10.7	644	10	BI435969
12	282.8	10.3	694	10	BI180003
13	278	10.2	572	10	BE345464
14	271.2	9.9	493	10	BE417348
15	268.4	9.8	368	9	AI065720
16	268	9.8	569	10	BE416603
17	262.4	9.6	588	10	BI922800

18	262	9.6	441	10	BE474181
19	258.4	9.4	504	9	AW906651
20	258.2	9.4	514	10	BE445315
21	256.4	9.4	596	9	AW060121
22	255.6	9.3	697	10	BI177125
23	251	9.2	559	10	BF728117
24	242.6	8.9	750	10	BI925772
25	237.4	8.7	412	10	BE415950
26	234.4	8.6	533	9	AW906977
27	232.8	8.5	753	10	BG593228
28	232.2	8.5	416	10	BI179933
29	222.2	8.1	360	10	BE416123
30	217.2	7.9	602	9	AW622870
31	216.2	7.9	401	10	BE417033
32	213	7.8	221	10	BE453927
33	212.8	7.8	640	10	BE129884
34	208.4	7.6	1084	9	BE035254
35	201.8	7.4	698	10	BI068228
36	198.4	7.3	408	10	BF052601
37	190.6	7.0	698	10	BI271813
38	186	6.8	660	10	BM259723
39	180.8	6.6	374	10	BG644021
40	173.4	6.3	439	9	AV532210
41	171.4	6.3	528	10	BE644323
42	171.4	6.3	607	10	BE517454
43	171.2	6.3	579	10	BI418178
44	170	6.2	577	9	AW720408
45	169.2	6.2	498	10	BM324234

ALIGNMENTS

RESULT 1

AW065905 500 bp mRNA linear EST 30-MAR-2000
LOCUS 687002G02.y1 687 - Early embryo from Delaware Zea mays cDNA, mRNA

ACCESSION AW065905

VERSION AW065905

KEYWORDS EST

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE 1 (bases 1 to 500)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

AUTHORS Walbot,V

TITLE Maize ESTs from various cDNA libraries sequenced at Stanford

JOURNAL Unpublished (1999)

COMMENT Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 687002 row: G column: 02.

Location/Qualifiers

1..500

source

features

features

features

features

features

features

features

features

features

features

features

features

features

features

features

features

features

features

features

features


```
QY 1793 gttgttatgtgtgccacagttcaggatgcgtgtgttcttctctctggtggtgtctttttcc 1842
||||| ||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 GGTGTCATGGGCTTCTCAGTTCAGAAATGGGGGTGTCTTTGGGATCTCTTCTCTACA 360
||||| ||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1843 atggtgtcctcaagtatctctctgccacctatgatgttgttagtcgtggaggagccct 1902
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 ATGGTCTCCCAAGATATCTCTCTACTCATGATGATGTGTGTCTGTGTGTCTCA 420
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1903 aatcccttcgcgaatgtatgttgcatatgacgacacacacacacacacacacacacac 1962
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 AATGCTTTTCGCAAGCATGTTCGATGATGACACATATAGCCCAATTTGAAGAAA 480
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1963 goctgttcttggcgatcaatttcagatttaccacaaagttcaaatcaccagcctaac 2022
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 GCTCTCCTTCTGTGATGCCATTTTCAGATCTGCCGAAGGTTGACAACATATCAACACAT 540
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2023 gatgtgatgagatgtgtgttcccccaagaccgaattccagcgtatcattcagctcagt 2082
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 GAGTGATGAATATCGTGTCAACCCCAAGACTGAATTCACGGGTATATACGACTTAGT 600
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2083 cgtaaagacatgttgatgtgttccttcctcctcgtgag 2116
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 CGTAAAGATATGTGGCATATTCATTGGTGATG 634
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 3

```
BF625060 841 bp mRNA linear EST 17-OCT-2001
LOCUS HVSMEa0005K18f Hordeum vulgare seedling shoot EST library
DEFINITION HVCDNA0001 (Cold stress) Hordeum vulgare cDNA clone HVSMEa0005K18f,
mRNA sequence.
ACCESSION BF625060
VERSION BF625060.2 GI:13081459
KEYWORDS EST.
SOURCE barley.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.
1 (bases 1 to 841)
Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Oates,R., Choi,D.W.
, Fenton,R.D. and Main,D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex cold-stressed seedling shoot cDNA
library
Unpublished (2001)
On Dec 18, 2000 this sequence version replaced gi:11888794.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 551
Seq primer: AATTAACTCTACTAAAGGG
High quality sequence stop: 817.
Location/Qualifiers
1..841
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEa0005K18f"
HVCdna0001 (Cold stress)"
/tissue.type="Seedling shoot"
/lab_host="TJCL21"
/note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
Seeds were surface sterilized then germinated under axenic
conditions in the dark at room temperature on filter paper
with water, nystatin and cefotaxime in covered
crystallization dishes. Five-day old seedlings were
```

FEATURES

source

```
incubated at 50C for 2 days. Shoots were then harvested,
total RNA was prepared, poly(A) RNA was purified, one
primary unamplified cDNA library was made, and 60000 pfu
in vivo excised to give pBluescript SK(-) cDNA
phagemids. These steps were performed in the TJ Close
laboratory at the University of California, Riverside
(Choi, Close, Fenton). Phagemids were plated and picked at
the Clemson University Genomics Institute (CUGI) (Begum,
Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations
DNA sequencing and sequence analysis were performed at
CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main
). The sequence has been trimmed to remove vector sequence
and contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders Also
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
BASE COUNT 190 a 201 c 251 g 199 t
ORIGIN
```

```
Query Match 14.9%; Score 407; DB 10; Length 841;
Best Local Similarity 74.0%; Pred. No. 1.8e-67;
Matches 529; Conservative 0; Mismatches 185; Indels 1; Gaps 1;
```

```
QY 2021 acgatgtgagtgatgtgttcccccaagaccgaattccagcgtatcattcgaactca 2080
||||| ||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ATGAGGTGATTTGAATATATCGTGTCAACCCCAAGACTGAATTCACGGGTATATACACTTA 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2081 gtcgtaaagacatgttggtgtctcctcgtgagggggtccagatgaaggaagc 2140
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GTCGTAAAGATATGTGGACTATTTCATTTGGTGATGACACTTGTCTGTAAGAAGCAAGC 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2141 tcttgatccagcgtttacagcttaacacgatgattagcgggttcaacagattc 2200
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 TCTTGGACCACAGCCTTTAAGCTAAACACGAGATGATTATGACCGAGTCCAGCAGATTC 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2201 ctgtcaagaaggagacaaacttcgcacctaaaggcgtgaggggttgagcaacaata 2260
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 COATAAAGAGGAGCAAAATTTCCGTGACTTCCAGCGCTCAAGGTTGAGCGAATACAA 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2261 ttgttgatgggagtcagaaatcgagcgtgtgaaactttcatctgggaaccactggttc 2320
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 TTGTGGAGTGGGATCCAGAAGTTGAGCGTGTCTACCTCAAGTCTGGCAAACTTTTGTGTC 300
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2321 ctgactatgcaatgtatcatcaagggaacaaatcaactcaacgcttttgccgttgtgt 2380
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 CTGACTATGCAATGTCTTCATCAAGGGCAGATCACCAGGCGTTTGCTCGTTGTGTGT 360
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2381 gggacgacagattcctacagttgtaaccagacagagcctcacaccagggttaataatc 2440
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 GGGATGAGACGGTGCCTACGTTGTACACAGACAGAGCCGCATACACAGATCATCTTCG 420
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2441 atccgactcaagaagggttcctcactatcccggaagacgcaagggttacagggtccccc 2500
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 ACCCGAACCGGAGCGTGTTCAGTGTCCGCGAGACGCTAGGCTGCAAGGGTTCGCCG 480
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2501 attactaccgattgttgcccgatcaaggagaagtacattcaagtcgggaacgcagtgg 2560
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 ACTACTACCGGATGACGGTTCCTCATGAAGGAGAAATATATCCAAAGTCGGAAGACGTGTGG 540
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2561 ctgtccctgtgcccgggcaactggcctactgtctggggcgcaagcctacctgggtgaactcg 2620
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 CTGTCCCGTCGCACGTGCACATGGGCTACTCTCTGGCGGCGGTACACAGGCGGAGGGTG 600
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2621 agggagtgaccctctgtaccagctgcctccaaagtttcaactctgttgaggagacgactg 2680
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 ATGCGGGCATGATGCGCTGTTTGTCTTCCCGAGAGCTTCACCAACGTCGCGGCGAGACCG 660
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Qy 2681 cggggcaggcgagggcctctct-gttggcaccctcgaggggaggtagttgagc 2734
| ||| |||| | | | |||| |||| ||| ||| ||| ||| ||| ||| |||
Db 661 GATCAGGGCAAGGGGCTCCTTCGTGGGACCTTGCCGGCTAGGTGGTCCATC 715

RESULT 4
BM073464/c
LOCUS MEST67-G10.T3 ISUM4-TN Zea mays cDNA clone MEST67-G10 3', mRNA
DEFINITION sequence.
ACCESSION BM073464
VERSION BM073464.1 GI:16917866
KEYWORDS EST.
SOURCE ze mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 717)
Qiu,F., Cui,F., Guo,L., Ashlock,D.A., Wen,T.J. and Schnable,P.S.
Expressed Sequence Tags from B73 Maize Seedlings and Silks
Unpublished (2001)
Contact: Patrick S. Schnable
Schnable Laboratory
Iowa State University
G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
Tel: 515-294-0975
Fax: 515-294-2299
Email: schnable@iastate.edu
Individual basecall and confidence value were assigned using the
Phred software,
(<http://depts.washington.edu/ventures/collabtr/direct/index.htm#b
rt>). Overall sequence quality assessment and vector trimming were
conducted using the Lucy software (<http://www.tigr.org/softlab/>).
Lucy parameters were set to ensure an overall trimmed quality of
97.5% or better without any vector fragments in the chosen
high-quality region of each sequence. Low-quality bases between the
poly-T and the high-quality region were replaced with N's to serve
as spacers.
PCR Primers
CTA TAG)
FORWARD: Forward PCR primer sequence, primer T7-1 (AA TAC GAC TCA
BACKWARD: Backward PCR primer sequence, primer T3 (ATT AAC CCT CAC
TAA AG)
Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).
FEATURES
Location/Qualifiers
1..717
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="MEST67-G10"
/clone_lib="ISUM4-TN"
/tissue_type="Seedling and silk"
/lab_host="DH10B"
/notes="Vector: pT73PAC; Site_1: EcoRI; Site_2: NotI;
ds-cDNA molecules were generated as follows. First-strand
cDNA was prepared from oligo-dT selected mRNA by priming
with a NotI oligo-dT primer (5'
AACTGGAAGATTCCGGCCGCGCAGGATTTTTTTTTTTTTTTT). The
resulting DNA:RNA hybrid was treated with RNase H and used
as a template for DNA PolI-catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the
ds-cDNAs were digested with NotI and size-selected. The
resulting molecules were directionally cloned into the
EcoRI and NotI sites of the pT73PAC vector. The library
then went through one round of normalization to Cor value
of 5 based on the methods of Marcelo Bento Soares (Genome
Research 6: 791-806, 1996)."
```

```
Best Local Similarity 99.7%; Pred. No. 4,1e-63;
Matches 384; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2352 atcaactcaagccgtttggcgccctgtgtgggacgagacagttcctacagttgtaaccag 2411
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 717 ATCATCAAGCCGTTTGGCGCCCTGTGTGGGACGAGACAGTTCNTACAGATTGTGTAACCG 658

Qy 2412 agcagagcctcaaacacagggttaataatccgactcaagcaagggtccctccactatccg 2471
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 657 AGCAGAGCCTCACACACAGGTTAATTCATCCGACTCAAGCAAGGTCCTCACTATCCG 598

Qy 2472 ggagaacgcaaggtttacagggcttccccgattactacagattgtttggcccgatcaagga 2531
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 597 GGAGAACGCAAGGTTTACAGGGCTTCCCCGATTACTACCGATTGTTTGGCCCGATCAAGGA 538

Qy 2532 gaagtacattcaagtcgggaacgagtcgctgtccctgttgcgcgggcactgggctactg 2591
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 537 GAAGTACATTCAAGTCGGGAACGCAGTGGCTGTCCCTGTGTCCCGGGCACTGGGCTACTG 478

Qy 2592 tctggggcgaagcctacctggtgaatctgaggggagtgacctctgtaccagctgcctcc 2651
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 477 TCTGGGGCAAGCCTACCTGGGTGAATCTGAGGGGAGTGACCCCTCTGTACCAAGCTGCCCTCC 418

Qy 2652 aagtttcacctctgttgaggacgacctgcgggcaggcgagggcctctctctgttgagcac 2711
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 417 AAGTTTCACTCTGTTGGAGGACGCACTGGCGGGCAGGCGGGGCTCTCTCTGTGGGCAC 358

Qy 2712 cctgcagggggaggtagttgagcag 2736
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 357 CCTGCGAGGGGAGGTAGTTGAGCAG 333

RESULT 5
BE345465 BE345465 558 bp mRNA linear EST 17-JUL-2000
LOCUS 946035B01.y1 946 - tassel primordium prepared by Schmidt lab Zea
DEFINITION mays cDNA, mRNA sequence.
ACCESSION BE345465
VERSION BE345465.1 GI:9254997
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 558)
REFERENCE Walbot,V.
AUTHORS Zea mays.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946035 row: B column: 01.
FEATURES
Location/Qualifiers
1..558
/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="946 - tassel primordium prepared by Schmidt
lab"
/tissue_type="tassels"
/dev_stage="Just after the transition from vegetative to
inflorescence development"
```

BASE COUNT	ORIGIN
118 a	143 c
170 g	127 t

Query Match 13.1%; Score 358; DB 10; Length 558;
Best Local Similarity 98.6%; Pred. No. 3.5e-58;
Matches 361; Conservative 0; Mismatches 5; Indels

QY	2371	gcctctggtggagagacagcttctacagtgttaacccagagcagcctccacaaccag	2430
Db	1	CGCCTGTGTGGGAGAGAGAGTTCCTACAGTGTAAACAGAGCAGAGCCTCAACACAG	60
QY	2431	gtataattcatccgactcaagcagggtctctcactatccggagagacgcaaggtttacag	2490
Db	61	GTATTATATTCATCCGACTCAAGCGAGGGTCTCTACTATCCGGGGAACGCAAGGTTTACAG	120
QY	2491	ggctcccgattactaccgattgttgcccatcaagagaaagatcattcaagtcggg	2550
Db	121	GCCTCCCTTGACTACTACCGATTGTTTGCCCGGATCAAGAGAGTACATTCAAGTTGG	180
QY	2551	aaccagtggtctccctgtgtcccgggcactgggtactctgtggggcaagcctacctg	2610
Db	181	AACGCAAGTGGCTGTCCCTGTGTCCCGAGCAGCTGGGGTACTGCTGGGGAAGCTACCTG	240
QY	2611	ggtgaatctcaggggagtgacctctgtaccagctgcctccaagtctcaacctctgttga	2670
Db	241	GTTGAATCTGAGGGAGTGACCTCTGTACCAAGTGCCTCCAAGTTTCACCTCTGTTGA	300
QY	2671	ggaagcaactcgggggcagcgaggcgctctcctctgtggcaacctcgaggggagtagtt	2730
Db	301	GGACGCACATGCGGGCAGGCGAGGGGCTCTCCTGTGGCACCCCTGCAGGGGAGTAGTT	360
QY	2731	gagcag	2736
Db	361	GAGCAG	366

RESULT	6
BI178113	
LOCUS	
DEFINITION	BI178113 761 bp mRNA linear EST 09-JUL-2001
ACCESSION	EST519058 cSTE Solanum tuberosum cDNA clone cSTE12D14 5' sequence, mRNA sequence.
VERSION	BI178113
	BI178113.1 GI:14643924

ORGANISM SOLANUM TUBEROSUM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterides I; Solanales; Solanaceae; Solanum.

1 (bases 1 to 761)

REFERENCE

AUTHORS van der Hoeven, R., Bezzlerides, J., Bachem, C., Visser, R., Cho, J., Chimento, A., Bougri, O., Buell, C. R., Ronning, C., Tanksley, S., and

TITLE Generation of ESTs from in vitro grown microtubers
 JOURNAL Unpublished (2001)
 COMMENT Contact: Cathy Ronning

CONTACT: Cathy Manning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13P-R

```

FEATURES
seq primer: 5'attc k...
            Location/Qualifiers
              1..761
                /organism="Solanum tuberosum"
                /cultivar="Bintje"
                /db_xref="taxon:4113"
                /clone="CSTE12D14"
                /clone_lib="cSTE"
                /tissue_type="axillary buds of stem explants; growing
                sink-tubers"
                /dev_stage="7, 8 and 10 days"
                /lab_host="SOLR"

```

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
 XhoI; Tissue supplied by Christian Bachem and Richard
 Visser (Department of Plant Breeding, Wageningen
 University, The Netherlands). The cSTA libraries will
 attempt to capture the induction and initiation/initial
 growth of the tuber in an in vitro system as described in
 Bachem et al. (Plant Journal, 1996). Small microtubers
 develop from axillary buds attached to stem explants when
 placed on a high sucrose medium (10%). Visible
 morphological changes occur synchronously at day five in
 the axillary buds. The first library, cSTA (1-20) consists
 of axillary buds harvested on days 1-3. This targets
 those genes involved in induction of the microtubers. The
 following libraries, cSTA (21-40) and cSTA (41-60),
 capture genes involved in tuber initiation and outgrowth.
 This library is noted as p3 in Tanksley lab notebooks."

[illegible]

```
RESULT 7
BE455437 590 bp mRNA linear EST 22-OCT-2001
LOCUS HVSMEG0017C03f Hordeum vulgare pre-anthesis spike EST library
DEFINITION HVCNDA0008 (white to yellow anther) Hordeum vulgare cDNA clone
HVSMEG0017C03f, mRNA sequence.
ACCESSION BE455437
VERSION BE455437.2 GI:13154758
KEYWORDS EST
SOURCE barley.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 590)
AUTHORS W. R. Close, T. J. Kleinjohs, A. Wise, R. Begum, D. Frisch, D. Yu,
Y. Henry, D. Palmer, M. Rambo, T. Simmons, J. Choi, D. W. Fenton
R. D. Close, S. J. Oates, R. and Main, D.
TITLE Development of a genetically and physically anchored EST resource
for barley genomics: Morex pre-anthesis spike cDNA library
JOURNAL Unpublished (2001)
COMMENT On Jul 26, 2000 this sequence version replaced gi:9465113.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 105
Seq primer: AATTACCCCTCACTAAAGGG
High quality sequence stop: 509.
Location/Qualifiers
1..590
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEG0017C03f"
/hvcdna_lib="Hordeum vulgare pre-anthesis spike EST library
HVCNDA0008 (white to yellow anther)"
/lab_host="SOLR"
/tissue_type="pre-anthesis spike"
/notes="vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
Plants were grown in the greenhouse at the University of
California, Riverside (Fenton, SJ Close, TJ Close). Whole
spike with awns trimmed were collected at white, green and
yellow anther stages (Fenton). Total RNA was prepared from
each pool, equal quantities of all three RNA pools were
combined, poly(A) RNA was purified from the mixture, one
primary unamplified cDNA library was made, and 1 million
pfu were in vivo excised to give pBluescript SK(-) cDNA
phagemids. These steps were performed in the TJ Close lab
(Choi) at the University of California, Riverside.
Phagemids were plated and picked at the Clemson University
Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins
and Wing) plasmid DNA preparations, DNA sequencing and
sequence analysis were performed at CUGI (Wing, Yu, Frisch
, Henry, Simmons, Oates, Rambo, Main). The sequence has
been trimmed to remove vector sequence and contains a
minimum of 100 bases of phred value 20 or above. For more
details on library preparation and sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders
Also see Close TJ, Wing R, Kleinjohs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)."
BASE COUNT 148 a 127 c 185 g 130 t
ORIGIN
```

Query Match 11.3%; Score 309.8; DB 10; Length 590;
Best Local Similarity 75.0%; Pred. No. 5.4e-49;

Matches	400;	Conservative	0;	Mismatches	132;	Indels	1;	Gaps	1;
QY 2074	cgactcagtcgtaagacacattgttgattggtcttcggtgaggggggttcgcagatgaa	2133							
Db 56	CGACTTTAGTCGGTAAGATATGTTGGACTATTTCATTTGGGTGATGACACTTGTCTGGAAGAA	115							
QY 2134	ggcaagctcttggtaccacagcctttacggcttaacaacagatgattatgagcgggttcaa	2193							
Db 116	GGCAAGCTCTTGGACCACCGCTTTAAGGCTTAACACGAGATGATTATGACCGGCTCCAG	175							
QY 2194	cagattcctgtcaagaaggagcacaacttcgcgcgaactaaaaggcgtaggggttgaaga	2253							
Db 176	CAGATTCGATTAAGAAGGGAGCATATTTGCGTGACTTCCCAAGCGTTAAAGTTGGAGCG	235							
QY 2254	aaacaattgttgagtgaggtccagaaatcgagcgtgtgaaacttctctgggaacc-	2312							
Db 236	AATAACATTTGGAGTGGAGTCCAGATGTTGAGCGTGTCTACCTCAAGTCTTGGCAGACCT	295							
QY 2313	actgggttcctgactatgcaatgtcattcatcaagggcacaatcaactcaagcgtttggcg	2372							
Db 296	TTTGGTCCCTGACTATGACAGTGTCTTCATCAAGGGCAGATCACCNAAGCGTTTGGTCG	355							
QY 2373	cctgtggtgggacgacagttcctcacagttgttaaccagagcagcctcaacaccaggt	2432							
Db 356	GTTGTGTGGATGAGAGCGGTGCTACGTTGTGACGAGCAGAGCGCATATCCAGAT	415							
QY 2433	tataattcatccgactcaagcaagggtcctcactatccggggagagcgaaggttacagg	2492							
Db 416	CATTCTGCACCGACAGGACGCTGTTTGTGACTGTCCGCGAGACGCTAGCTGCCAGG	475							
QY 2493	cttcgccgacttaccgattgtttggcccgatcaaggagagaagatcaagtcgggaa	2552							
Db 476	GTTCCCGCGACTACCTACCGGTGAACGGTTCATGAAGGAGAGTAGTACATCCNAAGTCGGAGA	535							
QY 2553	cgcagtggcttcctgtgtgccgggcactgggctactgtctcggggcgcaagcct	2605							
Db 536	CGCTGTGGCTGGTCCCGTCGCACGCTGCGCTTACTGTGTTGGCGGGCGT	588							
RESULT 8	AI586713/c								
LOCUS 486055E12.x4 486	- leaf primordia cDNA library from Hake lab Zea								
DEFINITION	mays cDNA, mRNA sequence.								
ACCESSION	AI586713								
VERSION	AI586713.1	GI:4573064							
KEYWORDS	EST.								
SOURCE	Zea mays.								
ORGANISM	Zea mays.								
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;								
AUTHORS	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC								
TITLE	clade; Panicoidae; Andropogoneae; Zea.								
JOURNAL	Walbot.V.								
COMMENT	Maize ESTs from various cDNA libraries sequenced at Stanford								
	University								
	Unpublished (1999)								
	Contact: Walbot V								
	Department of Biological Sciences								
	Stanford University								
	855 California Ave, Palo Alto, CA 94304, USA								
	Tel: 650 723 2227								
	Fax: 650 725 8221								
	Email: walbot@stanford.edu								
	Plate: 486055 row: E column: 12.								
FEATURES	Location/Qualifiers								
source	1..418								
	/organism="Zea mays"								
	/cultivar="B73"								
	/db_xref="taxon:4577"								
	/clone_lib="486 - leaf primordia cDNA library from Hake								
	lab"								
	/tissue_type="leaf primordia"								

```

/dev_stage="p7-pl1 leaf"
/lab_host="E.coli XLI-Blue MFR"
/note="Organ: shoot; Vector: Lambda zap; Hake lab cDNA library."
BASE COUNT      92 a   131 c   107 g   88 t
ORIGIN

```

```

Query Match      11.3%; Score 308; DB 9; Length 418;
Best Local Similarity 100.0%; Pred. No. 1.1e-48;
Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2429 aggtataattcatccgactcaagcaagggtcctcactatccgggagagcaagaggttac 2488
Db 418 AGGTTATAATTCATCCGACTCAAGCAAGGGTCTCACTATCCGGGAGAACGCAAGGTTAC 359
QY 2489 agggcttccccgattactaccgattgtttggcccgatcaagagaagtaacattcaagtcg 2548
Db 358 AGGGCTTCCCGGATTTACTACCGATTGTTGGCCCGATCAGGAGAGTAGTACATTCAGTCG 299
QY 2549 ggaacgcagtggtctccctgttggccgggcaactggggtactgtctggggcaagcctacc 2608
Db 298 GGAACGCAGTGGGTCTCCCTGTGGCCGGGCACTGGGCTACTGTCTGGGGCAAGCCTACC 239
QY 2609 tgggtgaatctgagggagtgacctctgtaccagctgctcctcaagtttcacctctgtg 2668
Db 238 TGGGTGAATCTGAGGGGAGTGACCTCTGTACACAGCTGGCTCCAGGTTTCACCTCTGTTG 179
QY 2669 gaggacgcactgcggggcagggcagggcctctctgttggcccccctgcagggaggtag 2728
Db 178 GAGGACGCAGTGGGGGCGAGGCGGGGCTCTCTGTGGTGGCACCCCTCGAGGGAGGTAG 119
QY 2729 ttgagcag 2736
Db 118 TTGAGCAG 111

```

```

RESULT 9
BE475748
LOCUS      BE475748      507 bp      mRNA      linear      EST 28-JUL-2000
DEFINITION 946048C03.Y2 946 - tassell primordium prepared by Schmidt lab Zea
ACCESSION  BE475748
VERSION     BE475748
KEYWORDS    EST.
SOURCE      Zea mays.
ORGANISM    Zea mays.
REFERENCE   1 (bases 1 to 507)
AUTHORS     Walbot,V.
TITLE       Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL     Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 946048 row: C column: 03.
FEATURES    Location/Qualifiers
             1..507
                /organism="Zea mays"
                /cultivar="OH43"
                /db_xref="taxon:4577"
                /clone_lib="946 - tassell primordium prepared by Schmidt lab"
                /tissue_type="tassels"
                /dev_stage="just after the transition from vegetative to inflorescence development"

```

```

/lab_host="XLRLR"
/note="Organ: tassels; Vector: HybriZAP; Site_1: EcoRI; Site_2: XhoI; George Chuck dissected immature tassels between 1mm and 3mm. Sharon Stanfield prepared the cDNA library in HybriZAP. Sample insert size range was 350 bp to 3 Kb with a 1 Kb average."
BASE COUNT      104 a   130 c   157 g   116 t
ORIGIN

```

```

Query Match      10.8%; Score 296; DB 10; Length 507;
Best Local Similarity 98.4%; Pred. No. 2.2e-46;
Matches 299; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2433 tataattcatccgactcaagcaagggtcctcactatccgggagagcaaggttacaggg 2492
Db 15 TATATTCATCCGACTCAAGCGAGGGTCTCTCATTATCCGGGAGAACGCAAGGTTACAGGG 74
QY 2493 cttccccgattactaccgattgtttggcccgatcaagagaagtaacattcaagtcggaa 2552
Db 75 CTTCCTGACTACTATCCGATTGTTGGCCCGATCAAGGAGAGTAGTATTCAGTTGGGAA 134
QY 2553 cgcagtggtctccctgttggccgggcaactggtctactgtctgtgggcaagcctacctggg 2612
Db 135 CGCAGTGGCTGTCTCTGTGGCCGAGCACTGGGCTACTGTCTGGGGCAAGCCTACTGGG 194
QY 2613 tgaatctgaaggagtagacctctgtaccagctgctcctcaagtttcacctctgttgagg 2672
Db 195 TGAATCTGAGGGGAGTAGACCTCTGTACAGGTGCTCCCAAGTTTCACCTCTGTGGAGG 254
QY 2673 acgcactgcggggcagggcagggcctctctgttggcccccctgcagggaggtagttga 2732
Db 255 ACGCACTGCGGGGCGAGGCGGCTCTCTCTGTGGCACCCCTGCAGGGAGGTAGTTGA 314
QY 2733 qcag 2736
Db 315 GCAG 318

```

```

RESULT 10
AW330561/c
LOCUS      AW330561/c      613 bp      mRNA      linear      EST 31-JAN-2000
DEFINITION 707029A05.x2 707 - Mixed adult tissues from Walbot lab (SK) Zea
ACCESSION  AW330561
VERSION     AW330561
KEYWORDS    EST.
SOURCE      Zea mays.
ORGANISM    Zea mays.
REFERENCE   1 (bases 1 to 613)
AUTHORS     Walbot,V.
TITLE       Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL     Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 707029 row: A column: 05.
FEATURES    Location/Qualifiers
             1..613
                /organism="Zea mays"
                /cultivar="W23"
                /db_xref="taxon:4577"
                /clone_lib="707 - Mixed adult tissues from Walbot lab (SK)"
                /tissue_type="tassel, kernel, silk, husk, root, leaf"

```

/dev_stage="adult"
/lab_host="PH10B"
/note="Organ: tassels, kernel, silk, husk, root, leaf:
vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."
BASE COUNT 164 a 176 c 148 g 125 t
ORIGIN

Query Match 10.8%; Score 295.4; DB 9; Length 613;
Best Local Similarity 98.0%; Pred. No. 3e-46; Mismatches 0; Gaps 0;
Matches 299; Conservative 0; Indels 6; Indels 0;
QY 2432 tiataatcaccgactcaagggctctcactatccgggagacgcaaggtttacagg 2491
Db 613 TTATAATTTCATCCGACTCAAGCGAGGGTCTCCTACTATCCGGGAGAACGCAAGGTTACAGG 554
QY 2492 gttcccccattactaccgatgtttggcccgatcaaggagaagatcaattcaagtcggga 2551
Db 553 GCTTCCCTGATTACTACCGATTGTTGGCCCGATCAAGGAGAAGTACATCAAGTCCGGGA 494
QY 2552 agcagtggtctgtccctgttggccgggcaactgggctactgtctggggcaagcctacctgg 2611
Db 493 ACGCAGTGGCTGTCCCTGTTGGCCGGGCACTAGGCTACTGTCTGGGGCAAGCCTACCTGG 434
QY 2612 gtgaatctgaaggagtagacctctgtaccagctgcctccaaagtttcaacctctgttgag 2671
Db 433 GTGAATCTGAGGGGAGTGACCCCTGTACAGCTGCCCTCCAAAGTTTCACCTCTGTGGGAG 374
QY 2672 gacgcactcggggcagggagggcctctctgttggcaccctcgacggggaggttagttg 2731
Db 373 GACGAACTCGGGGCAGCGAGGGCCTCTCCTGTTGGCACCAGCGAGGGAGGTAGTTG 314
QY 2732 agcag 2736
Db 313 ACAG 309

RESULT 11
BI435969
LOCUS EST538730 cSTD Solanum tuberosum cDNA clone cSTD21E9 5' sequence,
DEFINITION mRNA sequence.
ACCESSION BI435969
VERSION BI435969.1 GI:15260659
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 644)
AUTHORS van der Hoeven,R., Bezzerides,J., Fwing,E., Cho,J., Chiemingo,A.,
Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
TITLES Generations of ESTs from dormant potato tubers
JOURNAL Unpublished (2001)
COMMENT The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.

FEATURES
source Location/Qualifiers
1..644
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTD21E9"
/clone_lib="cSTD"
/tissue_type="dormant tuber"
/dev_stage="one month post-harvest"
/lab_host="SOLR"

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; This library targets genes expressed in dormant
tubers. This library was made from sections of dormant
tuber, avoiding the buds and epidermis. Tubers were stored
for one month post-harvest at 40C. The tuber was peeled
well away from the surface. Then it was chopped into 1-2
mm cubes and immediately frozen in liquid nitrogen. This
library is noted as P4 in Tanksley lab notebooks."
BASE COUNT 176 a 121 c 158 g 189 t
ORIGIN

Query Match 10.7%; Score 292; DB 10; Length 644;
Best Local Similarity 65.8%; Pred. No. 1.4e-45;
Matches 424; Conservative 0; Mismatches 220; Indels 0; Gaps 0;
QY 1429 gaaccgattgataaacttgactgactgccgcagaaaaatagagaattttgacaagaagg 1488
Db 1 GAGCCAATAGATGGCTTAAGTGACTGCCCGAAGAAATCAAAGAGTTTGTGGCGAAAGGC 60
QY 1489 cacaaagaagaattctcccactgcctggtgatgttgatgtcatttggaggagccacca 1548
Db 61 TTCAAAGCAAAATCTTTTCCCTTGCCCTGGGTGATGTGATGTGTATGTCGGGTGGACCACT 120
QY 1549 tgccaaggtatcagtggtgtttaaactgcgtacagaacacggtgatgcacactcaaatgatgag 1608
Db 121 TGCCAAGGAATCAGTGGATTCATATCGTTTTAGGAATAAAGAAATCCGATCGAGGATCCT 180
QY 1609 azaaacaacaagaatgtgactttcatgtatattgtggactgtgaagcccaagatgatt 1668
Db 181 AAAAATAAACAACTTGTATGTCTACATGGACATTTGGATTTCTTGAACCCAGCGTTTGTGA 240
QY 1669 ctcatggaaaaatgtgtggacatactcaaatcttgvggatggtttacctaggaaaaatgct 1728
Db 241 TTATATGAGAGACGTGGTGGACCTTGTCAAAATTTCTCAAAATGTTCTTGGSCGATATGCA 300
QY 1729 ttgaactgcttgtgtcatgaagaccagcgctgtgaatgatgtgtgctggtgtgc 1788
Db 301 TTGAGCAGACTTGTGGGATGACTACCAAGCACGATGGGAATGATGGCAGCTGTGTGCG 360
QY 1789 tatgtctgccacagttcaggatgctgtgttctctcgtgggtgtctcttcttctccatggtg 1848
Db 361 TATGGGTCTTCCACAATTTCTGATGTCGTCTTCATGTTTGGTGTCTTTTCATCAGAGAAG 420
QY 1849 ctccctaagtatctctctccacactgatgtgtagtacgtgaggagggcccttaatgccc 1908
Db 421 TTGCCACATATCCATTGCCACACATAAAGTTATTGTAAAGGGGTGTTATTCCCGTAGAA 480
QY 1909 ttttcgaatgtatggttgcatatgacgagacacaaaaacacatcccttgaaaaagccttg 1968
Db 481 TTTGAGTCAAAACACAGTTGCCGATATGATTACGCCCGGATCTTTGAGTTAAAGAAAGAACTC 540
QY 1969 ctctctggcgaatgcgaatttcagatttaccaaaagtttcaaaatcacaccagccttaacgatgtg 2028
Db 541 TTTCTTGGTGACGCACTTTCTGATCTCCCTTTGGGTGGAGACAATGAGCCCAAGAGATGAA 600
QY 2029 atggagatggtgtgttcccccaagaacgaattccacgctacat 2072
Db 601 ATGCTTTACACTGATGAGGCCCAATCTGATTTCCAGCATTTTAT 644

RESULT 12
BI180003
LOCUS EST520948 cSTE Solanum tuberosum cDNA clone cSTE21E10 5' sequence,
DEFINITION mRNA sequence.
ACCESSION BI180003
VERSION BI180003.1 GI:14645814
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

QY 2511 attgtttggccgatcaaggagaagtacattcaagtcgggaacgcagtggtctcctct 2570
|||||
Db 512 ATTGTTTGGCCCGATCAAGGAGAAGTACATTCAAAGTTGGGAACGAGTGGCTGTCCCTGT 453
|||||
QY 2571 tgcccgggcactgggtctactgtctggggccaagcctacctcgggtgaactctaggggagtga 2630
|||||
Db 452 TGCCGAGCACTGGGCTACTGTCTGGGGCAAGCCTACCTGGGTGAATCTGAGGGGAGTGA 393
|||||
QY 2631 ccctctgtaccagctgcctccaagtttcaacctctgttggaggagcagcactcggggcaggc 2690
|||||
Db 392 CCCCTGTCACCAAGCTGCCTCCAAAGTTTCACCTCTGTGTGAGAGCACTCGGGGGCAGGC 333
|||||
QY 2691 gagggcctctcctgttggcaccctcgagggggaggtagttgagcag 2736
|||||
Db 332 GAGGGCTCTCCTGTTGGCACCCCTCGAGGGGAGGTAGTTGAGCAG 287
|||||
RESULT 14
BE417348
LOCUS BE417348 493 bp mRNA linear EST 24-JUL-2000
DEFINITION MUG020.A11R990620 ITEC MUG Wheat Spikelet Library Triticum aestivum
CDNA clone MUG020.ALL, mRNA sequence.
ACCESSION BE417348
VERSION BE417348.1 GI:9415194
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 493)
AUTHORS Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
,S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,
Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,
Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,
Pecchioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,
Sorrells,M., Warburton,M. and Wenzel,G.
TITLE International Triticaceae EST Cooperative (ITEC): Production of
Expressed Sequence Tags for Species of the Triticeae
JOURNAL Unpublished (2000)
COMMENT Contact: Ogihara Y
Kihara Institute for Biological Research, Yokohama City University
Maioaka-cho 641-12, Totsuka-ku, Yokohama 244-0813, JAPAN
Tel: 81 45 820 1903
Fax: 81 45 820 1901
Email: ogihara@yokohama-cu.ac.jp
International Triticaceae EST Cooperative (ITEC)
http://wheat.pw.usda.gov/genome.
FEATURES
source
1..493
/organism="Triticum aestivum"
/cultivar="Norin 26"
/db_xref="taxon:4565"
/clone="MUG020.ALL"
/clone_lib="ITEC MUG Wheat Spikelet Library"
/tissue_type="young spikelets"
/dev_stage="Feekes" scale 6-7"
/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; M13 Reverse sequencing primer used. 1.2 kbp average
insert size."
BASE COUNT 103 a 125 c 154 g 111 t
ORIGIN
Query Match 9.9%; Score 271.2; DB 10; Length 493;
Best Local Similarity 74.3%; Pred. No. 1.2e-41;
Matches 342; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 2277 agaaatcagagcgtgtgaaatttcattctgggaacacactgttctactatgcaatgctc 2336
|||||
Db 1 AGAAGTTGAGCGGTGCTACCTCAAGTCCGGCAACCTTGTGTTCTCTGACTATGCAATGTC 60
|||||
QY 2337 attcatcagggaacaatcattcaagccgttttgggcgctgtggtgggacgagacagttcc 2396
|||||

Db 61 CTTATCAAGGCGAGATCACCAAAAGCCGTTTGTCGGTGTGTGGGATGAGACAGTTCC 120
|||||
QY 2397 tacagttgtaaccagagcagcagctcacaccagggttataattctccgactcaagcaag 2456
|||||
Db 121 CACGTCGTCGACAGACAGAGCGGCATACCAACAGATCATACTGCACCCGAATCAGGACG 180
|||||
QY 2457 ggtccctcactatccgggagaacgcaagttacagggtctcccgattactaccgattgtt 2516
|||||
Db 181 CGTTTTCAGCTGTCCTGTGAGAACAGAGCTTCAGGTTTTCCTCCGATTAATACCCGGATGTA 240
|||||
QY 2517 tggcccgatcaaggagaagtacattcaagtcgggaacgcagtggtctgctctgttgcgc 2576
|||||
Db 241 CGGTCCCATGAAGGAGAAGTACATCCAAAGTGGGAAACGGGTGTCTCTGTCCTGTGGCG 300
|||||
QY 2577 ggcactgggctactgtctggggcgaagcctacctcgtggtgaactcaggggagtgaccctct 2636
|||||
Db 301 GGCACGGGCTACTCTTTGGCCGCGGTACACAGGCGGAGGTGGGATGCGGTATGATGC 360
|||||
QY 2637 gtaccagctgcctccaagtttcaacctctgttggagacgcactcggggcaggcgagggc 2696
|||||
Db 361 GCTGTTTGTCTTCCTGTGACAGCTTCACCAACATCGGCACAGACCGGTGCGAGGGCAAGGGC 420
|||||
QY 2697 ctctcctgttggcaccctcgaggggaggttagttgagcag 2736
|||||
Db 421 CTCTCTCGTTGGCACCCCTCGAGCGGAGGTGTGTCGAGCAG 460
|||||
RESULT 15
AI065720/c
LOCUS AI065720 368 bp mRNA linear EST 24-JUL-1998
DEFINITION ag92b08.xl maize inflorescence immature ear library Zea mays cDNA
clone ag92b08 3', mRNA sequence.
ACCESSION AI065720
VERSION AI065720.1 GI:3341127
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 368)
AUTHORS Schutz,K., de la Bastide,M., Gnoj,L., Habermann,K., Huang,E.N.,
Parnell,L.D., Dedhia,N., Martienssen,R. and McCombie,W.R.
TITLE Expressed sequence tags from Z. mays
JOURNAL Unpublished (1998)
COMMENT Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: ag92 row: b column: 08
Seq primer: M13 forward universal -21
High quality sequence stop: 368.
FEATURES
source
1..368
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ag92b08"
/clone_lib="maize inflorescence immature ear library"
/sex="female"
/tissue_type="immature ear"
/note="vector: pBLUESCRIPT SK+ (X52325); Site_1: XhoI;
Site_2: EcoRI: This library is described in Schmidt, Hake,
et al.; (1993) Plant Cell 5:729-737. cDNAs are
directionally cloned into the XhoI and EcoRI sites; XhoI
is near the polyA tail. Most reads from this library are
3' in direction. Additional information on this library as
well as ftp access to all sequences can be found at
http://www.cshl.org/maizegenome"

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.
OM nucleic - nucleic search, using sw model
Run on: July 15, 2002, 08:28:03 ; Search time 4902.45 Seconds
(without alignments)
11678.855 Million cell updates/sec
Title: US-09-914-001-1
Perfect score: 2736
Sequence: 1 atggcgccgagctcccgctc.....cagggaggtagttgagcag 2736
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Query Match	Length	Description

1	2736	100.0	2796	8	AF243043
2	2264	82.7	3156	8	AY027539
3	450	16.4	2232	8	AF039365
4	445.8	16.3	2531	8	AF039364
c 5	260	9.5	17452	8	AC069324
c 6	216.8	7.9	127982	2	AC104473
c 7	174.6	6.4	85721	2	AC098835
8	173.6	6.3	410	8	AF433678
9	160.2	5.9	6605	8	AF383171
10	160.2	5.9	119111	8	ATF13C5
11	160.2	5.9	199075	8	ATCHRIV49
12	160.2	5.9	199382	8	ATCHRIV50
13	154.4	5.6	4999	8	AF364174
14	154.4	5.6	5025	8	AF383170
15	154.4	5.6	87148	8	AC013289
16	151.4	5.5	4364	8	AF039369
17	151.4	5.5	4364	8	AF039374
18	151.4	5.5	4595	8	AF039367
19	151.4	5.5	4614	8	AF039368
20	151.4	5.5	4896	8	AF039366
21	151.4	5.5	4920	8	AF039370
22	151.4	5.5	9642	8	AF039373
23	151.4	5.5	9643	8	AF039371
24	151.4	5.5	9643	8	AF039372
c 25	151.4	5.5	37570	8	ATU53501
c 26	151.4	5.5	109694	8	F23A5
27	149.2	5.5	296	8	AF433679
28	146.2	5.3	4400	8	AF039375
c 29	93.2	3.4	1931	6	AR083152
30	89.8	3.3	4720	8	ATHWETASE
c 31	88.6	3.2	318930	2	AC073495
32	88.2	3.2	84432	8	AC005850
c 33	88.2	3.2	87394	8	AC005882
34	88	3.2	186	8	AF047327
c 35	88	3.2	203984	2	AC067854
36	87.6	3.2	215004	2	AC020618
37	87.2	3.2	282611	2	AL645746
c 38	86.6	3.2	179149	9	AC009785
39	86.4	3.2	69070	2	AC099959
40	86	3.1	451	10	MUSIR3E2
c 41	85	3.1	235985	30	AC068013
c 42	84.8	3.1	13095	9	HSFGFR4G
c 43	84.4	3.1	206497	10	AC084389
c 44	84.4	3.1	219200	10	AL589701
45	84.2	3.1	202010	2	AC016961

ALIGNMENTS

RESULT 1	AF243043	2796 bp	linear	PLN 21-FEB-2001
LOCUS	Zea mays DNA cytosine methyltransferase MET2a (met2a) mRNA,			
DEFINITION	complete cds.			
ACCESSION	AF243043			
VERSION	AF243043.1	GI:13021689		
KEYWORDS	Zea mays.			
SOURCE	Zea mays.			
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.			
REFERENCE	1 (bases 1 to 2796)			
AUTHORS	Papa,C.M., Muszynski,M.G., Springer,N.M. and Kaeppler,S.M.			
TITLE	Maize chromomethylase, Zea Methyltransferase2a (zmet2a), is required for CpNpg Methylation			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 2796)			
AUTHORS	Papa,C.M., Muszynski,M.G., Springer,N.M. and Kaeppler,S.M.			
TITLE	Direct Submission			
JOURNAL	Submitted (08-MAR-2000) Agronomy and Plant Genetics, University of Minnesota, 1991 Buford Circle, St. Paul, MN 55108, USA			

FEATURES
source

Location/Qualifiers
1..2796
/organism="Zea mays"
/db_xref="taxon:4577"
/chromosome="10"
1..2796
/note="genotype: Mol17"
/gene="met2a"
31..2769
/gene="met2a"
/note="met2a; chromomethylase"
/codon_start=1
/product="DNA cytosine methyltransferase MET2a"
/protein_id="AAK11516.1"
/db_xref="GI:13021690"
/translation="MAPSSPSPAAPTRVSRKRAAKAEIHKNEEEVEEVAASAKR
SRKAASGKKPKSPKQAKPGKKKDAEMKEPEVDDVCARPEDEELAMGEERAEPO
AMOEYVAVANGSPGKKRVGRNAAAGDHEPFTGSPVADEARSWPKRYGRSTA
AKPDEEBELKARCHRSKAYDNVNYCLGDDVYKAGENEADYIGRTIEFEFGTDQCH
YFCRMFRAEDTVINSLSVSDGHKHDPRRVFLSEKNNDVLDCLISKKYIVHDP
NMDPKAQLIESDLYYDMSYSVAYSTFANISSENGSGSDTASGSDSDVDLEYS
SMPTRTATLDLYSGCGMSTGLCLGAALSGLKLETRWAVDFNSFACOSLYKNHPTE
VRNEKADEFLLAKWEAVLCRKYQVDYDLSNASEDADESPDKDEFVYELVIGIC
YGGSDRENGIYFKVQWEGYEPEDTWEPIDNLSDCPQKIREFVGEHKRKLPLPGDV
DVICGGPPCOGIGSFENRNDEPLKDEKNQWTPMDIVAYLKPVYLMENVVDILK
FADYLGKVALSCLVAMKYQARLGMVAGCYGLPQPMRVFLWAGALSMLVLPKPLPT
YDVYVRGAPNAPFSOCWAYDETQKPSLKKALLLGDALISDLPKVONHQPNDVMEYGS
PKTEFYQYIRLSKMDLWSFEGAGPDEGKLLDHQPLRLNDDYERVOQIIPVKKGAN
FRDLKGVYRGANNVHWPDEIERVLSKSGRLPVPDYMFSFKIGKSLRPFGRLMWDET
PTVYRAEPHNQVVIHPTQARVLITRENARLQGFDPDYRLFGPIKEKYLQVGNVAVP
VARALGYCLGQAYLGESGSDPLQLPSPSTSYGGRTAGQARASVPVGTGPAVEVVEQ"
BASE COUNT 701 a 645 c 820 g 630 t
ORIGIN

Query Match 100.0%; Score 2736; DB 8; Length 2796;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 atggcgagctcccggtcaccgcgcgcctacacgctctctggtggcgagcgcc 60
Db 31 ATGGCGCGAGCTCCGCTACCGCGCGCGCTACACGCGTCTTGGCGCGAAGCGCGC 90
Qy 61 gccaggcgagagatccaccgaacaggagagagagaggtcgcgcgcgctc 120
Db 91 GCCAAGCGGAGGATCCACGACAGAGAGAGAGAGAGGTGCGCGCGGCGTCC 150
Qy 121 tccgcaagcgagcgcaaggcgcatcttcggggaagagcccaagtcgcccccaag 180
Db 151 TCGGCCAAGCGCAGCGCGCGCATCTTCGGGGAAGAGGCCAAGTCGCCCGCCCAAG 210
Qy 181 caggcccaagcgaggaagaagaggggatgccgagatgaaggagccgtggagagc 240
Db 211 CAGGCCAAGCGGGAGGAGAGAGAGGGGATGCCGAGATGAAGAGCCCGTGGAGGAC 270
Qy 241 gacgtgtcgcgaggagcccgacgaggaggattggccatggcgaggaggagcgag 300
Db 271 GACGTGTGCGCGAGAGCCGACGAGAGAGAGTGGCCATGGCGAGGAGGCGCGAG 330
Qy 301 gacgagccatgcagagagaggtggttcggttcgcgcggggtcacccgggaagaagag 360
Db 331 GAGCAGGCCATCAGGAGGAGGTGTTGCGGTGCGCGGGGTCAACCGGGAGAGAGG 390
Qy 361 gtggggaagaagcgcgcgcgcgctgtgcgaccagcgaggtcatcgtcgagc 420
Db 391 GTGGGAGAAGAAGACGCGCGCGCGCTGCGACACGAGCGGGAGTTTCATCGCGAGC 450
Qy 421 cctgttcgcgcgagcgcgcgcaactggcccaagcgctacgcgcgcgagcactgcc 480
Db 451 CTTGTTGCGCGGAGCGGCGCAGCAACTGGCCCAAGCGCTACGCGCGCAGCACTGCC 510
Qy 481 gcaagaacgcgatgagggaagagctcaaggccagatgtcaatcgcgagcgctaa 540
|||||

Db 511 GCAAGAAACCCGGATGAGGAGGAAGAGCTCAAGGCCAGATGTCACTACCGGAGCGCTAAG 570
Qy 541 gtggacaacgctcgtctactcgtcctcgggatgacgctctatgtcaaggctggagaaaaagag 600
Db 571 GTGGACAACGCTGCTACTTGCCTCGGGATGACGCTATGTCAAGGCTGGAGAAAAGAG 630
Qy 601 qcagattacattggcgcatcactactgaattttttgaggagactgacagtgctcattttt 660
Db 631 CGAGATTACATTGGCGCATTTACTGAATTTTTTGGAGGACTGACCAGTGTCACTATTTT 690
Qy 661 acttgccgtgtgttcttcgagagcagagacacgggttatcaattctttgtgttcataagt 720
Db 691 ACTTGCCTGTGTCTTCCGAGCAGAGGACACGGTATCAATTCCTTTGGTGTCCATTAAGT 750
Qy 721 gtggatggccacaagcagtcagccctagacgctgtttcttcttgagagaaaaaacgacaat 780
Db 751 GTGGATGGCCACAAGCATGACCCCTAGACGTGTTTCTTCTTGAGGAAAAGACGACAAT 810
Qy 781 gtgcttgattgatctatctcacaaggtcaagatagtcacatgttgatccaaaatatggatcca 840
Db 811 GTGCTTGATTGCTATATCTCCAAAGGTCAAGATAGTCCATGTTGATGCCAAATATGGATCCA 870
Qy 841 aaagcaagcgctcagctgatagagagttgcgaacctatactatgacatgtcttactctgtt 900
Db 871 AAAGCCAAGGCTCAGCTGATAGAGAGTTGCGACCTATACTATGACATGCTCTTACTCTGTT 930
Qy 901 gcatacttcacatttgcttaatactctctgaaatgggcaagtcaagcagtgatcacgct 960
Db 931 GCATATTCACATTTGCTTAATATCTCGCTGAAAATGGGCACTAGGCGAGTGATACCCT 990
Qy 961 tcgggtattctctctgatgagtgtggaactgtggagactcatctagtatgccaacgagaca 1020
Db 991 TCGGGTATTCTTCTGATGATGTGGATCTGGAGACGTCATCTAGTATGCCAACGAGACA 1050
Qy 1021 gcaacctcttgatctgtctctgctgtggggcatctctactggtcttctgttggt 1080
Db 1051 GCAACCTCTTCTGATCTGTATTCCTGCTGTGGGGCATGCTCTACTGTGCTTGGTGGGT 1110
Qy 1081 gcagctcttctggttgaactgaaactgaaactcgatgggctgtgtgtgttcaacagtttcg 1140
Db 1111 GCAGCTCTTCTTGGCTTGAAACTTTGAAACTCGATGGGCTGTGTGATTTTCAACAGTTTTGCG 1170
Qy 1141 tgcaaatgttaaaataaatccacagactgaggtgcgaaaatgagaaagccgagag 1200
Db 1171 TGCCAAAGTTTAAATATAATCATCATCACAGACTGAGTGCAGAAATGACAAAGCCGATGAG 1230
Qy 1201 ttcttgccctctcttaaggaatgggcaagttctctatgcaaaaaatatgtccaaagatgtggt 1260
Db 1231 TTTCTTGGCCCTCTTAAAGAAATGGGCGAGTTCTATGCAAAAAATATGTCCAAAGATGTGGAT 1290
Qy 1261 tcaaaatttagcaagctcagagatcaagcggatgaagacagagccctcttgcagaagcgaa 1320
Db 1291 TCAAAATTTAGCAAGCTCAGAGGATCAAGCGGATGAAGACAGACCCCTCTTTGACAAGGACGAA 1350
Qy 1321 ttgttgtagagaagcttgtcgggatatgttatggtgagtgacagagaaaaatggcatc 1380
Db 1351 TTTTGTGTAGAGAAGCTTGTTCGGGATATGTTATGTTGGTGGCAGTGACAGGAAAATGGGATC 1410
Qy 1381 tattttaaggtccagtggggaaggtacggccctgagggatcacatggtggaaccgattgat 1440
Db 1411 TATTTTAAAGTCCAGTGGGAAGGATACGGCCCTGAGGAGGATACATGGGAACCGATTGAT 1470
Qy 1441 aacttgagtgactgccccagaaaaattagagaatttgtacaagaagggccacaaagaag 1500
Db 1471 AACTTGAGTGACTGCCCGCAAAATTTAGAGAAATTTGTACAAAGAGGGCCACAAAGAAG 1530
Qy 1501 attctccactgcctggtgtgattgtgtatgtctattgttgaggccccaccatgccagggtatc 1560
Db 1531 ATTCTCCCACTGCCTGGTGTGATGTCTATTGTTGGAGGCCACCATGTCGAAGGTATC 1590
Qy 1561 agtgggtttaatcgtgcagaaaacgctgtagagccactcaagatgagaaaaacaaca 1620
Db 1591 ASTGGGTTTAAATCGGTACAGAAACCGGTGATGACCCACTCAANAGATGACAAAAACAACA 1650

Qy	1621	atgggtgactttcatgatattgtgcgtaacttaagacccaagfatgttctcatgaaaaat	1680
Db	1651	ATGGGTGACTTTCAFGGATATTGTGGCGTACTGTGAAGCCCAAGFATGTCTCATGAAAAAT	1710
Qy	1681	gtgggtgacatactcaaaatttcggatggttacctagagaaaaatgctttagctgcctt	1740
Db	1711	GTGGTGGACATACTCAAATTTGGGATGGTTACCTAGAAAAATATGCTTTGAGCTGCCTT	1770
Qy	1741	gttgcctaaagtaccaaagcgcttggaatgatgtgctggttgctataggctcaca	1800
Db	1771	GTTGCTATGAAGTACCAACGCGGGCTTGGAAATGATGGTGCTGGTGTGATGGTCTGCCA	1830
Qy	1801	cagttcaggatgcgtgttctctgtgggtgctcttcttccatgggtgcctcctaagtat	1860
Db	1831	CAGTTTCAGATGCGTGTGTCTCTGGGGTGTCTTTCTTCTCCATGTGCTGCCCTAAGTAT	1890
Qy	1861	ctcttgccacactatgatgttagtcgtgagagcccttaagtcctttctcgaatgt	1920
Db	1891	CCTCTGCCACCTATGATGTTGTAGTACGTGGAGAGCCCTTAATGCCCTTTCGCAATGT	1950
Qy	1921	atggttgatatgacgagacacaaaacatccctgaaaaagccttgcttttgccgat	1980
Db	1951	ATGGTGTGATATGACGAGACACAAAACCATCCCTGAAAAAGCCTTGCTTCTTTGGCGAT	2010
Qy	1981	gcaatttcagatttaccaaaggttcaaaatcaccagccttaacgatgtgatggataggt	2040
Db	2011	GCAATTCAGATTACCAAAGGTTCAAAATCACCAGCCTTAACGATGTGATGGATATGGT	2070
Qy	2041	ggttcccccaagccgaattccagcgctacatttcgactcagtcgtaagacatgttgat	2100
Db	2071	GGTTCCCCCAAGACCGGAATTCAGCGCTACATTCGACTCGTCTGAAGACATGTTGGAT	2130
Qy	2101	tggtccttcggtgaggggctgtccagatgaaggcaagctcttggatcaccagccttta	2160
Db	2131	TGSTCTTCGTGTGAGGGGCTGTCCAGATGAAGGCAAGCTCTTGATCACCAGCCTTGA	2190
Qy	2161	cgggttaacacgatgatattagcggggttcaacagattcctgtcaagaagggagccaac	2220
Db	2191	CGGCTTAACAACGATGATTATGAGCGGGTTCAACAGATTCTCTCAAGAAGGAGGCCAAC	2250
Qy	2221	ttccgcgacctaaaggcgctgaggggttgagcaacaatatgttgtagtgggataccagaa	2280
Db	2251	TTTCCGCGACCTTAAGGGCGTGAAGGTTGGAGCAACAATAATTGTTGAGTGGATCCAGAA	2310
Qy	2281	atcagcgtgtgaacttcaatctgggaaccactggttctgaactatgcaatgtcattc	2340
Db	2311	ATCGAGCGTGTGAACCTTTCATCTGGGAACCACTGGTTCCTGACTATGCAATGTCATTC	2370
Qy	2341	ataaggggcaaatcaactcaagccgttttggggccctgtgtggtggacgagacagttcctaca	2400
Db	2371	ATCAAGGGCAAAATCACTCAAGCGCTTTTGGGGCCCTGTGGTGGACGAGACAGTTCTCTACA	2430
Qy	2401	gttgaacacagacagagcctcaacacaggttataattcatccgactcaagaagggtc	2460
Db	2431	GTTGTAAACACGACGAGAGCCTCACACCAGGTATATTATCCGACTCAACACAGGGTC	2490
Qy	2461	ctcaactatccgggagaaacgaagggttacagggtctccccgattactaccogattgttggc	2520
Db	2491	CTCACTATCCGGGAGAACCAAGGTTACAGGGCTTCCCCGATTACTACGATTTGTTGGC	2550
Qy	2521	ccgatacgaaggagaagtacattcaagtccgggaacgcagtggtctcctgttgcgcgggca	2580
Db	2551	CCGATCAAGGAGAAGTACATTCAAGTCGGGAACGACGTGGCTGTCCCTGTTCGCCGGGCA	2610
Qy	2581	ctgggctactgtctgggggaagcctacctgggtggaattctagggggagtgacctctgtac	2640
Db	2611	CTGGGCTACTGTCTGGGGCAAGCCTACCTGGGTGAATCTGAGGGGAGTGACCCCTCTGTAC	2670
Qy	2641	cagctgcctccaagtctcaacctgtgtgagagacgcactcgggggcagggcagggccctct	2700
Db	2671	CAGCTGCCTCCAAGTTTCACCTCTGTGTGAGGAGACGCACTGCGGGGCGGAGGCGCCCT	2730
Qy	2701	cctgttgaccacctgcagggagagtaattgaacag	2736
Db	2731	CCTGTGGCACCCCTGCAGGGGAGGTAGTTGAGCAG	2766
RESULT	2		
AY027539			
LOCUS	AY027539	3156 bp	mRNA linear PLN 10-MAR-2001
DEFINITION	zeta mays chromomethylase (Zmet5) mRNA, complete cds.		
ACCESSION	AY027539		
VERSION	AY027539.1	GI:13272198	
KEYWORDS	zeta mays.		
SOURCE	zeta mays.		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.		
REFERENCE	1 (bases 1 to 3156)		
AUTHORS	Papa,C.M., Springer,N.M., Muszynski,M.G. and Kaeppler,S.M.		
TITLE	Maize chromomethylase Zea Methyltransferase 2 (Zmet2) is required for CpnpG Methylation		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 3156)		
AUTHORS	Springer,N.M., Papa,C.M. and Kaeppler,S.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-FEB-2001) Agronomy, University of Wisconsin, 1575 Linden Dr, Madison, WI 53706, USA		
FEATURES	Location/Qualifiers		
source	1..3156		
	/organism="Zea mays"		
	/db_xref="taxon:4577"		
	/chromosome="2"		
	/note="genotype: Mol17"		
gene	1..3156		
	/gene="Zmet5"		
CDS	57..2804		
	/gene="Zmet5"		
	/note="ZMET5"		
	/codon_start=1		
	/product="chromomethylase"		
	/protein_id="AAK15805.1"		
	/db_xref="GI:13272198"		
	/translation="MAPSPSARPTRASGRSAMAEEIHONBEEEEEVVAATAKRRRAASSGKKPKPTPOAKPAVAGKKGETKTEPVDVDCABEPDEELAMGEEEAEEQAMQEVAAVAAGSPGKRVGRSAASGDHVPFVIGSPVGAFAEAMNPKRYERSTAANKPREDDELKARCHYSKVDNIVYICLGDVVYKAGENADYIGRTTFEFGTDRCHYTCRFFRAEDTVINSLVNDVGHKHPRVFVSEKNDNDVLDCLISVKVLYHVPNDMPKAKQLTEHCOLDLYDMSYATSTFANTISENGSGSETASGISDDAGLETSSNMPERTATLLDLYSGCGMSTGLCLGAALSGLKLETRWVDLNSFACQSLKYNHPTQEVREKAEDEFLALLKENAVLCEKYVHQVDNSLAGEDEQDLDKDEFVQKLIGIRDGTGRKKGVTFKYQWEYGEEDTWEPIDNLSDCLPKIRFVQEGRKKILPLPGDVIDCGPPCQGISGFNFRNRDEPLDKENKOMTMDIVAYLKPVKVLMENVVDILKFDAGLYGLKVALSCLVAMKYQARGLMMVAGCYGLQFRMRVFLMGALSVMVLPKYPLTYDVVYRGGAAPNAFSQCVAYDETQPSLKKALLGLDPSLDLPKVHNHPNDVMEYGGSPKTEQRYIRLGRKMDLWDSFGEAGPDEGLKHDPRLRNDDDYERVKOIPVKKGANFDLKGKVGANNVVDPEVRVYLSGKPLVPDYAMSTFKCKSLKPFGLRWQDQVPTVTVRAEPHNQVILHETQARVLTIRENARLQCFDPDYRLFGPIKEKVIQVGNAAVAVPARALYCLGQALYGSDESQDPLQPLPASFTSVGRVAVQANAASVGTGPAGEVVEQ"		
BASE COUNT	813 a	686 c	885 g 772 t
ORIGIN			
Query Match	82.7%;	Score 2264;	DB 8; Length 3156;
Best Local Similarity	90.5%;	Pred. No. 0;	
Matches 2492;	Conservative	0; Mismatches	235; Indels 27; Gaps 6;
Qy	1	atggcgcgagctcccgctcacccgcgcgcgtacacgcgtctctctggggagcgccgc	60
Db	57	ATGGCGCCAAAGTCCCGCTCATCGCTAGCCACGCGCTCCGCGGGAGCGCTCC	116
Qy	61	gccaaaggccgagagatccaccagaacaagagagagaggttcgagcgcgctcc	120
Db			

Db 117 GCCATGGCGGAGAGATCCACCAGAAACCAGAGAGGAGGAGGAGTGTGTGGCTGCGTCC 176
Qy 121 tcgcgaagcgcagcgaagcgcgcatcttcgggaagaagcccaagtcgcgcccccaag 180
Db 177 ACCGCCACGCCAGACGCAAGCGCGCTCTCCGGGAAGAAGCCCAAGCGAGCCCAAG 236
Qy 181 caggccaa-----gccgggggaagaagggggatgcccagatgaagagagccc 231
Db 237 CAGGCCAAACCGCGCGTGGCGGGGATGAAGAAGAAGGGGGAGACCGAAGACGAGCC 296
Qy 232 gTggagagcagctgtcgccgagagagccgacgagagaggttgccatgggccaagga- 290
Db 297 GTTGTGGACGCTGTGCGCCGAGGAGCCCGACGAGGAGAGTTGGCCATGGCGGAGAG 356
Qy 291 -----ggaggccgagagcagccatgcyagggagggtgggttgccgctcgccggggta 345
Db 357 GAGGCGGAGGCGGAGGAGCGCATGCGAGGAGTGGTGGCGCGGTGCGCGCGGGTGG 416
Qy 346 ccgcggaagaaggtggggaagaagcgcgcgcgcgcgtgcgacacacagagccg 405
Db 417 CCGGTAAGAGAGGGTGGGAGAAGGAG--CGCCGCCGCTCGGGCGACCACTGTGCC 473
Qy 406 gagttoatcgagccctgttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 465
Db 474 GAGTTTCATCGGAGCCCGCTGTGTGGCGCGGCGGCGACAGCAACTGGCCCAAGCGCTAC 533
Qy 466 ggcgcgacactgcgcgaagaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 525
Db 534 GAGCGCAGCACTGCGGCANATAAACCGGAGGAGGAGCATGAGCTCAAGGCCAGGCGCCAC 593
Qy 526 taocggagcgctaaggtggacaacgtcgctactgcctcgcgcgcgcgcgcgcgcgcgcgc 585
Db 594 TACCGGAGCGCTAAGGTGGACAATGTCTACTCGCTCGGGGATGACGTTTATGTCAA 653
Qy 586 gctggagaagagcagagatcacattggccgcattactgaatttttgaggagactgac 645
Db 654 CTGGAGAAATGAGGCTGATTACATTTGTTCGCATTACTGAATTTTGGAGGCACTGAT 713
Qy 646 cagtgtoactatttaacttgcgctgggttcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 705
Db 714 CGATGCCACTATTTTACTTGGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 773
Qy 706 ttggttcccaaaagtgtgagtgcccaagcatgacccctagacgtgttttcttcttcgag 765
Db 774 TTGGTGTGATTAATGTGTGTCGCCAACAGCATGACCCCTAGACGTGTTTTCTTCCGAG 833
Qy 766 gaaaagaacgacatgtgcttgattgcattatctccaaggtccaagatgagtcgaattgat 825
Db 834 GAAAAGATGACAATGTGCTTGATTCATCTCCAAGGTCAAGATAGTCCATGTTGAT 893
Qy 826 ccaaaataggaTccaaaagccaaggtcgagctgagagagttgagccctatactatgac 885
Db 894 CCAATATGATTCCAAAAGCTAAGGCCCAACTGATAGAACAATTTGACCTTACTATGAC 953
Qy 886 atgtctactctgttgatattctacatttgcataatactcgtcgtaaaatgggcaagtc 945
Db 954 ATGCTCTACTGTGTGCATATTCATATTGTCTAATATCTCATCAGAAAATGGGCACTCA 1013
Qy 946 ggcagTgaTaccgcttcgggtatttctctctgagtgagtgagctggagagcgtcatctagt 1005
Db 1014 GGCAGTGAAACCGCTTCGGGTATTCTTCTGATGATGCGGGTTTGGAGACGTCATCTAAT 1073
Qy 1006 atgcaacgaggaagaacaccttcttgatctgtattctcgtgtgagggcagtgctact 1065
Db 1074 ATGCCAGAGAGGACAGCAACCCCTCTGTGACTTGTATTCGGGCTGTGGGGCATGTCCACT 1133
Qy 1066 ggtcttctgttggttgagctcttcttgcgttgaaacttgaaactcgatgggctgtgat 1125
Db 1134 GGTCTTTGCTTGGGTGCTGCTTCTTCTTGGCCTGAAACTTGAACATCGATGGGCTGTGAT 1193
Qy 1126 tcaaacagtttgccggtccaaagtttaaaataataatcatcaacagactgaggtgcgaat 1185
Db 1194 CTTAACAGATTTTGGTGGCCAAAGTTTAAATAATCAATATCATCCACAGACTGAGGTGCCAAAT 1253

Qy 1186 gagaagccgatgagtttcttgcctcctctaaagaaattgggcagttcttatgcacaaaatat 1245
Db 1254 GAGAAAGCTGATGAGTTTCTGCTCTCCTTAAGGAATGGCAGTCTTATGTCAAAAATAT 1313
Qy 1246 gt---ccaagatgtggattcacaatttagcaagctcagaggtacaaagcgagatgaagacagc 1302
Db 1314 GTTCATCAGACGTGGATTCGAATTTAGCAGGTTTCAGAGGATCAGAGGATGCTGACAA-- 1371
Qy 1303 cctctgaacaaggaagaaattgtttagaagacttgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1362
Db 1372 -CTCTTGACAAAGGATGAATTTCTGTACAGAAGCTTATTTGGAATACGCTATGATGGCACT 1430
Qy 1363 gacaggaaaaagtgcatctattttaagggtccagtggaaggtacagccctgagagagat 1422
Db 1431 GGAAGGAAAAAGGCGCTATATTTAAGGTCCNATGGGAAGAAATATGGTCTCTGAAGAGGAT 1490
Qy 1423 acatgggaacccgatgataacttgagtgactgcccgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1482
Db 1491 ACATGGGAACCTATTGATAACCTGAGTGACTGTCCGCTGAAAAATTAGAGAAATTTGTACAA 1550
Qy 1483 gaaggcacaagaagaagatctccactgcctggtgatgtatgtatgtatgtatgtatgtatgtatgt 1542
Db 1551 GAAGGGCGCAAAAAGAAAAATCTCCCGCTTCTGGTGATGTGTGATGTCTATTTGTGGAGGC 1610
Qy 1543 ccacatgccaaaggtatcagtggtgttaatacgcgtacgaacacgcgtgatgagccactcaaa 1602
Db 1611 CCACCATGTCAAGGTATAAGTGGGTTAATCGGTTCAAGGAACCGTGATGAGCACTTAAA 1670
Qy 1603 gatgaaaaaacaacaatggtgacttctatgatatgtatgtatgtatgtatgtatgtatgtatgtatgt 1662
Db 1671 GATGAGAAAAAACGAATAGTGTGACTTTCTATGATATTTGTGCATATTTGAGGCCCAAG 1730
Qy 1663 tatgtctcatggaaaaatgtgtggaca tactcaaaatttgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1722
Db 1731 TATGTTCTCATGAAAAATGCTGGACATACTGAAATTTGCTGGCTACCTAGGAAAA 1790
Qy 1723 tatccttgagctgccttctgtctatgaagtaccaaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1782
Db 1791 TATGCTTTGAGCTGCTCTAGTTGCTATGAAGTACCACGACGCTTGGAAATGATGTTGGCA 1850
Qy 1783 ggttgctatggttcgcacagttcaggatgcgtgtgttctctctggtggtgcttctctctcc 1842
Db 1851 GGCTGCTATGCTGTCACAGTTCAGGATGCGTGTGTCTCTGGGGTCTCTTCTTCTCC 1910
Qy 1843 atgctgctcctcaagtatcctctgcccacctatgattgttagtcagtgagagagccct 1902
Db 1911 ATGCTGCTCCCTAAGTATCCTCTGCTACATATGACGTTGTAGTACGTGGAGGACACCT 1970
Qy 1903 aatgccttttcgcaatgtatgtgtgcatatgacgcgagacacaaaaaccatccctgaaaaaaa 1962
Db 1971 AATGCCCTTTTCGCAATGTATGTTGCTCATATGATGAGACACAAAAGGCCATCTCCCTGAAGAAA 2030
Qy 1963 gcctgctcttggcgatgcaatttcagatttaccagggttcaaaatccacagcctaac 2022
Db 2031 CTTTGGCTTCTTGGTGATGCTGATTTTACAAAGTTCGAAAACCATCAACCTAAC 2090
Qy 2023 gatgtgatggagtatgtggtgttcccccaagaccgaaattccagcgctacattcgcactcagt 2082
Db 2091 GATGTAATGGAGTATGGTGGTTCGCCCAAGACAGAGTTCCAGCGCTACATTCGACTTGGT 2150
Qy 2083 cgtaaagacatgttggatgtgtccttcggtgaggggctggtccagatgaagcagctc 2142
Db 2151 CGTAAACATATTTGGATTTGTGCTTTGCTGAGGAGGCTGGTCCAGATGAAGCAAGCTC 2210
Qy 2143 ttgatcaccagcctttcaggttaacaacgatattatgagcgggttcaacagattcct 2202
Db 2211 TTGGATCACCAGCCCTTACGGCTTAACAATGATGATTTATGAGCGGGTTTAAAGCAATTCCT 2270
Qy 2203 gTcaagagggagcacaactctcgacacctaagcgctgaggggttggagcaacaataatt 2262
Db 2271 GTCAGAAGGGAGGCCAACCTTCCGTGACCTTAAAGGCTGTCAAGGTTGGAGCAATTAATGTT 2330

QY 1250 -----aagatgtgattcaaaatttagcaagctcagagatcaagcgatgaa 1296
Db 856 ATTTCAAGCACTGAACCAATGAATCAATTC TGAGTTGGAAGTGAACAAAGTGACGAA 915
QY 1297 gaagccctcttgacaag-----gaogaatttggtagta 1329
Db 916 AATGATGACATAGATGAGGCAAGTACTGATGGAACCTCTCAGCTGGAGAGTTTGAAGTA 975
QY 1330 gagaagcttgcgggatatgttatgtgtgcagtg-----cagggaaaatggcatctat 1383
Db 976 GAAAGTTTGTGGAAATCGTTTTTGGAGATCCTTAAGGCACAGGGGAAAGACACTTCAT 1035
QY 1384 tttaaggtccagtyggaagatacggccctgagggagatacatatgggaaccgattgataac 1443
Db 1036 CTTAAGGTAAAGGTGAAGGATACAGTCTTAAGTATGATACATGGGAACCTTACTCTGGC 1095
QY 1444 ttgagtactgcccgcagaaaaattagaaatttgtacaagaaggccacaaaaagaagatt 1503
Db 1096 TTAGGCAATTTGTAAGGAGAAAGCTCAAGGAGTATGTGACAGATGGGTTCAAGTCTCATCTT 1155
QY 1504 ctcccactgctggtgatgttatgtatcatttgtgagggccaccatgccaaggtatcagt 1563
Db 1156 TTACCACTTCCCGGTACTGTTTACTCTGTGTGGGTGGGCCACCCCTGCCAAGGGATCAGT 1215
QY 1564 ggggttaatcgttacagaaacgctgatgagccactcaagatgagaaaaacaaacaaatg 1623
Db 1216 GGTATTAACCGGTTTAGAACAACAACACGCCGCTTGAGGATTAAGAAAACACGACTT 1275
QY 1624 gtgacttcatgatattggtggtcacttgaagcccaagtatgtctcatggaaaaatgtg 1683
Db 1276 TTGTGTGTTTTAGACATCATTGACTTCTCAAGCCAAAGTTTACGTGCTTATGSAAAATGTT 1335
QY 1684 gtggacatactcaaaatttgcggatggttacctaggaataatgctttgagctgccttggt 1743
Db 1336 GTAGACCTTCTTAGGTTTCAAAGGGTTACCTGCGCGTCATGCTGTGCTAGCTTGT 1395
QY 1744 gctatgaagtacaaagcggccttggaatgatggtggctggttgctatggtctgccacag 1803
Db 1396 GCCATGAATTTACCAGACGCGGTTGGCGATGATGACTGCTGGATTCCTACGGTCTCCCTCAG 1455
QY 1804 ttcaggatgcgtgtgtctctcgtgggtgctcttcttcccatggtgcctcctaagtatcct 1863
Db 1456 GTTAGAACAGGGTTTTTTTATGGCGCGCACAAACACAGAGAAATCCGCCCTTTACCCA 1515
QY 1864 ctgccaactatgatgttagtagacgtggaagagcccttaagtccttttcgcaatgatg 1923
Db 1516 CTTCCACACATGAAGTTGAGAAAAATPACAAACACACCAAGGAATTTGAGGATTTACAA 1575
QY 1924 gttgcatatgacgagacacaaaaaccatccctgaaaaagccttgcctcttggcgatgca 1983
Db 1576 GTTGGACTTATTCAAAAGGAGCTTCTTCAATTAGACAACGCTCTTACGCTGGCTGATGCT 1635
QY 1984 atttcagatttcaaaaggttcaaaataccacagcctaagatgtagtgagt---atggt 2040
Db 1636 ATTAGTGACCTACCCCGGTGACAAATTTCTGAGGCAAAATGATGTAAAGGAATTAATGAT 1695
QY 2041 ggttccccaagaccgaattccagcgctacatttcgactcagtcgtaaaagacatgttgat 2100
Db 1696 GCTGCTCCTTAAGACAGATTTTGAGAACTTTATTAGCCTTA-----AAAGATCTGAG 1746
QY 2101 tggctctcgtgaggggctggtccagatgaaggcaagctcttgatcacacagccttta 2160
Db 1747 ACATATTACCCTATGTGTGGAGATCCGGCAGCGAGGTTGTTTGATCATCACCCCTCT 1806
QY 2161 cggcttaacacgatgattatgagcgggttcaacagattcctgtcgaagaaggggagccaac 2220
Db 1807 GAATTAAGAGATGACGACTTGGAAAGAATTTCCCTACATACCAAAAAAAGAGGGGCAAA 1866
QY 2221 ttccgcacctaagaaggcgtgaggttgagcaacaataatgttgtgagtggaatccagaa 2280
Db 1867 TTTGCTGATATGCTGCTGTTTATGTT---CACAAACAAAGCCCACTTAACCTTAGG 1923
QY 2281 atcagcgtgtgaaactttcatctgtgggaaccactggttctgactatgcaatgcatc 2340

Db 1924 GTGAAGCGAGCAAAAGTTAAAAATCTGGAAGAATGTGGTACCTGATTATGCAGTTTCATTC 1983
QY 2341 atcaagggccaatacactcaagccgcttittgggcctctgtgtgtggagcagagatctctaca 2400
Db 1984 ATAAAGGAAAAATCAAAAAGCCCTTTTGGTAGATTGTGTGGGACGAATAAGTTTAAACG 2043
QY 2401 gttgtaaccagagcagagcctcacaaccagggttataatctccgactccaagcaagggtc 2460
Db 2044 GTTGTCACAAGGGCAGACCTCACACCCAGTGGGTAAATCCATCCAAATGCAAGACGCTG 2103
QY 2461 ctcaatctccggggaagcagcaaggtttacagggcttcccogattactaccgatgtttg 2520
Db 2104 TTGTCTGTCCGTGAGAAATGCCAGACTACAAGGTTTCCCTGATTGTGTAAAGCTTTGTG 2163
QY 2521 ccgatcaagagagaatcattcaagtcgggaacgcagtggtctccctgttgcgcgggca 2580
Db 2164 TCAATTAAAGAAAGTACATCCAGGTGGGAATGCAGTGGCAGTTCTCTGTAGGGGTGCG 2223
QY 2581 ctgggcta 2588
Db 2224 TTAGGATA 2231

RESULT 4
AF039364 Arabidopsis thaliana ecotype kl-0 chromomethylase (CMT1) mRNA, complete cds.
LOCUS AF039364
DEFINITION AF039364.1 GI:2766712
ACCESSION AF039364
VERSION AF039364.1
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 2531)
AUTHORS Henikoff,S. and Comal,L.
TITLE A DNA methyltransferase homolog with a chromodomain exists in multiple polymorphic forms in Arabidopsis
JOURNAL Genetics 149 (1), 307-318 (1998)
MEDLINE 98250692
PUBMED 9584105
REFERENCE 2 (bases 1 to 2531)
AUTHORS Henikoff,S. and Comal,L.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1997) HHMI, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N (A1-162), Seattle, WA 98109-1024, USA
FEATURES
source
1. .2531
/organism="Arabidopsis thaliana"
/strain="ecotype kl-0"
/db_xref="taxon:3702"
/chromosome="1"
1. .2531
/gene="CMT1"
/allele="kl-0"
1. .2376
/gene="CMT1"
/codon_start=1
/product="chromomethylase"
/protein_id="AAB93485.1"
/db_xref="gi:2766713"
/translation="MAARNKQKRAEPESDLCFAGKPMVVESTIRWPHRYQSKTKLQVAPTKPKANGKKDEETIKQAKHFDKALVDGLINLNDVVYVTGPGKLFIAKVELFADDDGVPYCRFRWYRPEDTLIERFSLVQPKRVFLSNDENDNPLTCTWSKVNIAKVLPTKTSRIEQRVIPPQYDDMKVEYPLNFI SADDSADSSSSSSSALNCFE NLHKDKFLDLYSGCGAMSTGFCMGASISGVKLITKWSVDINKFACDSLKNHPETE VRNEADEFLALKWKRLCEKESLVSTBPVESISELDEFEVENDIDDEASTGAEL EPEFEFEVFLGMEFGDFQGTGKGTQLQMLRWKGYNSSDTWEPYSGLGNCKELKEI VIDGFKSHLLPLUPGTIVTYCCGPPCGISGTNRYRNEAPLEDKNQQLLVFLDILDF LKPNYVLMNVNDLLRFKSGFLARHVAASFAMNYOTRIGLMMAGSYGLPOLNRVFL

WAAQSEKLPPLPDPHEVAKKFNPNPKFKDLQVGRIOEMFLKLDNALTLADALSILP
PVTNYVANDYMDINDAAKPTBFENFISLKRSETLLPACGDPTRLFDHQLVLGDDDD
LERVSYIPKOGANTYDMPEGLVHNHNAEINPRFRALKSGKNVVPYAFATSFYKGSK
KPFGRWGBDEIVNTVTRAEPHNQCVJHPMNRVLSVRENARLQGFDPCKLQGTIKE
KYIQGVNAVAVPVGVALGYAFGMAOGLTDEDPVILKLPKPYPECMQANDQI"
BASE COUNT 759 a 496 c 583 g 693 t
ORIGIN

Query Match	16.3%	Score 445.8	DB 8	Length 2531
Best Local Similarity	54.5%	Pred. No. 2.1e-64		
Matches 1177	Conservative 0	Mismatches 877	Indels 105	Gaps 10
Qy 484	aagaaacggatgaggaagagctcaaggccagatgtcactcaccgagcgctaaagggtg	543		
Db 169	AAAAAAGAAGATGAAGAGATTATAAAGCAAGCTAAATGCCACTTTGACAAAGCTTTGGTT	228		
Qy 544	gacaacgtctactcctcgggagatgacgtctatgtcaaggctggagaaacaggaca	603		
Db 229	GATGGAGTTCATCAACCTAAACGACGATGCTATGTCACGGGTCTACCGGGAAGCTG	288		
Qy 604	gattacattggccgcatctactgaattttttgaggggactgaccagtgctactattttact	663		
Db 289	AAATTTCATTGCAAAAGTTATCGAGCTTTTGAAGCTGATGATGGGGTCCCGTATTGTGCG	348		
Qy 664	tgccgttggttctccgagcagagacacggttatcaattcttgggtgccataagtg	723		
Db 349	TTTCGATGTTATTCGGCTGAAGTACTTTAATCGAAGATT-----	392		
Qy 724	gatggccacagcatgacccctagacgtgtttttcttcttgggaagaacacacatgtg	783		
Db 393	--TTCCCACTTGTGCAACCAAGAGAGTGTCTCTCAAAATGATGAGATGACAAATCCT	450		
Qy 784	cttgattgcatctatctccaaaggtcaagatagtcctatgt---tgatccaaatattgatcca	840		
Db 451	CTCACCTGTCATCTGTCAAAAGTCAACATCGCTAAAGTCCCATTTACCTAAGATTACCTCC	510		
Qy 841	aaagcgaagctcagctgatagagadtgcgacctatctatgacatgtctactctgtt	900		
Db 511	AGAAATCGCAAGGGTATTCTCTCTGTGATTTACTACTAGATATGAAGTAC-----	564		
Qy 901	gsatattacatttgcataatctcgtctgaaatgggcaagtcagcagtgataccgct	960		
Db 565	-----GAAGTTCCTTATCTCAACTTTATAAGTCGAGATGATGATGATGATGATGATGATG	615		
Qy 961	tcgggtatttctctgatgtggtgctgagacgtcatctatgattgccaacaggagaca	1020		
Db 616	TCAAGTCTTTCAAGTGACTCT---GCCTTGAATTCCTTTGAATCTCCACAGGATGAA	672		
Qy 1021	gaacccctcttgatctgtattctggctggtggggcagtgctactggtcttctggtgggt	1080		
Db 673	AAATTCTTATGGATCTATACAGTGGGTGCGAGCGATGTCACAGGTTTCTGCAATGGGA	732		
Qy 1081	gaagctcttctctggttgaacttgaactcgatggctgtgtgatttcaacagttttgctg	1140		
Db 733	GCTTCATATCAGGAGTAAATTTGATTAGCAAAATGCTGTGGATATAAATTAAGTTGCT	792		
Qy 1141	tgccaaagttaaataataatcatccacagactgaggtgcgaaatgagaaacccgatgag	1200		
Db 793	TGTGATAGTCTTAAACTTAAATCATCTGAAACAGAGGTTGAGAAACCCGCAAGAT	852		
Qy 1201	ttcttgccctcttaaggatggcagttctctatgcaaaaatat-----	1245		
Db 853	TTTCTAGCTCTCTCAAGAATGAAAAGGCTCTGTGAGAAATTTTCTCTCGTTTCAAGC	912		
Qy 1246	gtccaaagtgtgattcaaatattagcaagctcagagatcaacgcatgaagacagccct	1305		
Db 913	ACCGAACAGTGGAAATCAATTTCTGATTAGAAAGATGAGAGGTTGAAGAAATGATGAC	972		
Qy 1306	cttgacaag-----gacgaatttggtagaagott	1338		
Db 973	ATAGATGAGGCAAGTACTGGTGTGAACCTCGAAGCTTGAGAGATTGGAAGTAGAAAAGTTC	1032		

Qy 1339	gtcgggatatgttatgggtggcagtg-----cagggaataatggcatctattttaaggctc	1392		
Db 1033	CTTTGGCATCATGTTTGGAGATCTCAAGGCACAGGGGAAAGACACACTTCAACTTATGGTA	1092		
Qy 1393	cagtggaaggaatcagccctcggaggagatcacatggaacagcagattgataactgaatgac	1452		
Db 1093	AGGTGAAGGGATACAACTCTAGTTATGACACATGGAACCTTACTCTGGCTTAGGCACAC	1152		
Qy 1453	tgcccgacagaaatbttagaagaatttgtaagaaggccacaaagaaagattctcccaactg	1512		
Db 1153	TGCAAGGAGAACTGAAGGAGTACGTATAGATGGGTTTAAAGTCTCATCTCTTACCACCTT	1212		
Qy 1513	ccgtggtgatgttgatgtcatcttgaggagcccaactgccaaggttatcagtggtgttaaat	1572		
Db 1213	CCGGTACTGCTACACAGTGTGGGTGGGCCCTTGGCAAGGATCAGTGGGTATAC	1272		
Qy 1573	cggtagaagaaacccgtgatgagccactcaagatgagaaagaaacaaacaaatgggtgaacttc	1632		
Db 1273	CGGTATAGAAACAATGAAGCACCCGCTTGAAGATCGAGAGAACCCAGCAACTTTTGGTGT	1332		
Qy 1633	atggatattgtggctacttgaagcccaagtatgtctctcatggaaaaatgtggtggacata	1692		
Db 1333	TTAGACATCATTTGACTTCTCAAGCCAAACTATGTCTGATGGAAAAATGTTGTAGACCTT	1392		
Qy 1693	ctcaaatgttgaggtgttaacctaggaataatgtcttgagcgtccctgtgtgtatgaag	1752		
Db 1393	CTTAGGTTTTCAAAGGGTTCTCTGGCACGTCATGCTGTAGCTAGTCTTTGTGGGATGAAT	1452		
Qy 1753	taccaagcggcttgggaatgatgtggtggtggtctgtgctgccacagttcaggtg	1812		
Db 1453	TACACAGACGGTGTGGGAATGATGGCTGTGGATCTTACGGGCTCCCTCAGCTTAGAAAC	1512		
Qy 1813	cgtgttctctcgtgggtgctcttcttccatggtgctccctaaagtatcctcgtgccacc	1872		
Db 1513	AGGGTTTTTATGGCTGTCACACCATCAGAGAACTTCCCTCTTATCCCTCCCAACA	1572		
Qy 1873	tatgatgtgtgtagtcgtggaaggagccctaatgctcttcttgcagtgtatggtgtgcatat	1932		
Db 1573	CATGAAGTTGCTTAAAGAAATTCACACACCTTAAAGAAATTCAAAGATTTTACAAGTTGACGT	1632		
Qy 1933	gacgagacacaaacccatccctgaaagagcgtgctcttcttggcgatgcaatttcagat	1992		
Db 1633	ATTCAAAATGGAGTCTCTTAAATTTGACACACGCTCTTACGCTGGCTGATGCTATTAGTGAT	1692		
Qy 1993	ttaccaaaagttaaataatccacagcctaagatgtgagaggt---atggtgttcccc	2049		
Db 1693	CTGCCACCGGTGACAAATATGTGGCAATGATGTGATGATGATTAATGATGTGCTCTCT	1752		
Qy 2050	aagaccgaattccagcgtacatctcagtcagtcgtaagacatgttggattggtccttc	2109		
Db 1753	AAGACAGAGTTTGAAGACTTTTATTTAGCTCAAAAGATCTGAGACACTG-----TTA	1803		
Qy 2110	ggtagagggtggttccagatgaagcagcctcttgatccaccagccttaccggttaac	2169		
Db 1804	CCGCGATGTGTGGAGATCCGACACGAGGTTTGTGTGATCACCACCCCTCGTATTAGGA	1863		
Qy 2170	aacgatgattatgagcgggttcaacagattcctctgaagaaagggagcgaactcccgac	2229		
Db 1864	GATGATGACTTGGAAAGAGTTTCTCATACCAAAAACAGAGGGGGCAAACTATCCGCAC	1923		
Qy 2230	ctaaagggtgaggttggagcaacaattgttgagtgggtgggtccagaaatcgagcgt	2289		
Db 1924	ATGCTTG-----GTGTTTTTATTGTTCAACAACAAGCGGAAATTAATCTAGTGTTCGA	1977		
Qy 2290	gtgaacttctcatctggaaacacactggttctctgactatgaatgtcattcatcaagggc	2349		
Db 1978	GCAAGTTGAAATCTGGAATAAATGTGGTCCCTCTGCTTTATGCAATTTTCATTTAAAGGA	2037		
Qy 2350	aatcactcaagccgtttgggcgctgtggtggagcagacagttctcactagttgaac	2409		
Db 2038	AAATCAAAAAAGCCGTTTGGTAGACTTTGGGGTGTGAAATCGTTAAACACTGTTGTGACA	2097		
Qy 2410	agagcagagcctcacaaccagggttataattcattccgactcagcgaagggtcctcactatc	2469		

/protein_id="AAK00415.2"
/db_xref="GI:18542926"
/translation="MRAAVPPERKRSERRSAPWPTTLVAWQHSYPELRQSLPLP
KENLAKYLRWREKLPRGRLPPADYSYIKIGDVCEENTAVLSLWELGASEPACVY
YGDVAGAGADLLSLGRLSFGGAARHITVSPS"
complement(28339..28478)
/note="Similar to GatJin_Os3 MITE element"
complement(29362..30896)
/gene="OSJNBa0071K19.5"
/note="Contains similarity to homeobox protein HAT22"
complement(join(29362..29671,30181..30260,30543..30896))
/gene="OSJNBa0071K19.5"
/codon_start=1
/product="Putative homeobox protein HAT22"
/protein_id="AAK00416.1"
/db_xref="GI:12643022"
/translation="MAQDDGVGLALGSLGSGHRRQRESRDPASSAASLTLRL
PAESGQPVVVKREYVRAEEYEYERALYSSSAAAADDGCRKKLRSLKQ
SALLEDFKEHSLNPKQKVALAKOLNLRQVEVWFONRRARLKLKQTEYDCELLR
CCETLENNRRLHRELQOLRALHSTAAGFFMATLTPVPAATLSICPSCERLATAAA
GASPTAAADRNTKPTAPHLFSFPAKSAAC"
30919..30936
/note="Simple Sequence Repeat (CT)n"
31259..31280
/note="Simple Sequence Repeat (AT)n"
32569..32622
/note="Simple Sequence Repeat (TA)n"
33091..33332
/note="Similar to Tourist_O12 MITE element"
37648..37689
/note="Simple Sequence Repeat (AT)n"
42438..42455
/note="Simple Sequence Repeat (AT)n"
42453..42496
/note="Simple Sequence Repeat (TATG)n"
complement(46188..51797)
/gene="OSJNBa0071K19.6"
/note="Contains similarity to inositol 1,3,4-trisphosphate
5/6-kinase"
complement(join(46188..46282,46816..46900,46974..47029,
47466..47642,47744..47863,48062..48154,48606..48746,
48830..48862,49059..49140,51658..51797))
/gene="OSJNBa0071K19.6"
/codon_start=1
/product="Putative inositol 1,3,4-trisphosphate
5/6-kinase"
/protein_id="AAK00417.2"
/db_xref="GI:18542927"
/translation="MPLSPPLTAATAAVVAVAGQRLVGVYALTKKKVSLQPKLLS
LARKSHFVSDIETRLSEQGPDILHLKLTDEKQCVLEDTREHFEVIVLPFPA
IOLHNRQSMQLQEVADNLNSAYGEVCTPQOLVIMKDPISPSAVAKAGLTPLVAKP
LVDPGTSKSHLSLAYVETSLMDPLVLQEFVNHGGILFKYVVGETIRVVRFL
PDVNIYDLENDGIFRPVRSVCATAEADVDPSIAELFPKPLEKIGRELRRLGL
RLFNFDIREHGRKDRYYVIDINYPGKMGPGYEHIFIDFLLSLVQNKYKRRLLSGS"
51513..51536
/note="Simple Sequence Repeat (AATA)n"
57301..57362
/note="Simple Sequence Repeat (AT)n"
57915..58027
/note="Similar to Stowaway type MITE element"
complement(58688..59026)
/note="Similar to Tourist type MITE element"
59590..59646
/note="Simple Sequence Repeat (TAT)n"
complement(65305..65478)
/gene="OSJNBa0071K19.7"
/note="Hypothetical protein"
complement(65305..65478)

Query Match 9.5%; Score 260; DB 8; Length 174752;
Best Local Similarity 86.4%; Pred. No. 1.5e-33;
Matches 287; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1516 ggtatgtgtatgtcatcttggaggccaccatgccaagtgatcagtggtggttaatcgg 1575
Db 126519 GGTATGTTGATGTCAATTTGTTGGCGCACCATGCAAGAAATAAGTGTTCATCGG 126460
QY 1576 tacagaacgctatagaccactaaatgatgagaaacaaacaaatgtagcttctcatg 1635
Db 126459 TTCAGAACCCCAAGAGACCACTAAAGACGAGAAACAAACAGATGGTACATTCATG 126400
QY 1636 gatattgtggctacttgagcccaagtattctctcatggtgaaatgtggtgacatactc 1695
Db 126399 GATATTGTGCATACCTGAAGCTAAGTATGCTCTGATGCTGCAAAATGTTAGACATACTG 126340
QY 1696 aaatttcgcatgttacctagagaaatattctttagctgccttattctatgaagtac 1755
Db 126339 AAGTTTGCATATGGTTACCTTGACGCTTATGCTTGACCGCTGTGGGCTATGAAGTAC 126280
QY 1756 caagcgcggttggaatgatggtggtggttctatggttgcacagttcaggtatcggt 1815
Db 126279 CAAGCAGCAGTTGGAATGATGTTGAGGTTTGTATGCTTGCCACAATTTAGGATCGGT 126220
QY 1816 gtgttcctctggggtgactcttcttctccatggt 1847
Db 126219 GTGTTCTCTGGGGAGGCTCTCTTACCATGGT 126188
RESULT 6
AC104473/c
LOCUS
DEFINITION Oryza sativa chromosome 3 clone OJ126B05, *** SEQUENCING IN
PROGRESS ***, 7 ordered pieces.
ACCESSION AC104473
VERSION AC104473.1 GI:17530688
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
AUTHORS
Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H.-R., Rambo,T.,
Saski,C., Henry,D., Oates,R. and Simmons,J.
TITLE
JOURNAL
REFERENCE
AUTHORS
Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H.-R., Rambo,T.,
Saski,C., Henry,D., Oates,R. and Simmons,J.
TITLE
JOURNAL
COMMENT
Submitted (12-DEC-2001) Clemson University Genomics Institute,
Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 9763: contig of 9763 bp in length
* 9764 18964: contig of 9201 bp in length
* 18965 40158: contig of 21194 bp in length
* 40159 45784: contig of 5626 bp in length
* 45785 49464: contig of 3680 bp in length
* 49465 57772: contig of 8308 bp in length
* 57773 127982: contig of 70210 bp in length.
The nucleotide sequence of this BAC clone was generated by
combining Monsanto and Clemson University Genomics Institute

LOCUS	ATF13C5	119111 bp	DNA	linear	PLN 10-MAR-2000
DEFINITION	Arabidopsis thaliana DNA chromosome 4, BAC clone F13C5 (ESSA project).				
ACCESSION	AL021711				
VERSION	AL021711.2	GI:5738363			
KEYWORDS	thale cress.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
	1 (bases 4495 to 119111)				
	Bevan,M., Pohl,T., Weizenegger,T., Bancroft,I., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 4778)				
AUTHORS	Bevan,M., Robben,J., Grymonprez,B., Volckaert,G., Bancroft,I., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.				
JOURNAL	Unpublished				
REFERENCE	3 (bases 1 to 119111)				
AUTHORS	EU Arabidopsis sequencing,project.				
TITLE	Direct Submission				
JOURNAL	Submitted (09-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk				
COMMENT	On Aug 18, 1999 this sequence version replaced gi:2832611. Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.				
FEATURES	Location/Qualifiers				
source	1..119111				
	/organism="Arabidopsis thaliana"				
	/variety="Columbia"				
	/db_xref="taxon:3702"				
	/chromosome="4"				
misc_feature	1..4494				
	/note="positions 1-4494 are not included within BAC clone F13C5"				
exon	2221..3267				
	/gene="AT4g18830"				
	/number=1				
gene	2221..3267				
	/gene="AT4g18830"				
CDS	2221..3267				
	/gene="AT4g18830"				
	/note="similarity to MICROTUBULE ASSOCIATED PROTEIN 1A (FRAGMENT) - HOMO SAPIENS, EMBL:O95643"				
	/codon_start=1				
	/product="putative protein"				
	/protein_id="CA52868.1"				
	/db_xref="GI:5738364"				
	/translation="MRWGRKPVSSSSGLSRALPVSFWSKLSGGSDLKPAKEKKQD EKASNTSVKTSSTTRSDTHSKRFQVSEKENSATRSADRKEKEFEETMS VRKVRDFQKTCGFELEAMDNDGTVLPRIQVNRDQRCERDDQLLEQPKPRS KQAGVKVPARRTGGYSREDSVILGHTITKPAHWEKLEKLVREVLKDAQOOR KSLYLKRLNRIGTKNNKRVFSPRASEKCRKAIEDLKKAKORAREHELLIETADG GWNESEAVKCSSDQKDFRDSMIEMINGINHPEELKELLYCVLRLTWTDYHDMI TSVFQVHNDENFH"				
gene	3488..5125				
	/gene="AT4g18840"				
exon	complement(3488..5125)				
	/gene="AT4g18840"				
	/number=1				
gene	complement(3488..5125)				
	/gene="AT4g18840"				
CDS	complement(3488..5125)				
	/note="similarity to hypothetical proteins - Arabidopsis thaliana				
	Contains Prokaryotic membrane lipoprotein lipid attachment				

site AA304-314:Prokaryotic membrane lipoprotein lipid attachment site AA405-415"					
/codon_start=1					
/product="putative protein"					
/protein_id="CAAL6741.2"					
/db_xref="GI:5738365"					
/db_xref="SPTREMBL:O49399"					
/translation="MVKDIYMCACIIIRPOAYNLRLLQENLKKMSVCSSTPVPILS FTERAKSLTEIOAHAFMLKTGLFHDTSASKLVAFAANTPEKTVSAHSLILNRIGS PNGFTNSVIRAYANSTPEVALTVFREMLLGPVPDKYSFTFLKACAACFGEEGR QIUGLIGSLVDFVVENTLVNYGRSGIFEIARKYLDRMFVRDAYSWNLSLLSAYILE KGLVDARALFDEMEERNVSNWMISSYAAAGLVKREAREVEDSPMDYVYVSNWAVY AYAHVCQYNEVLVFNKMLDDSTPDGFTLVSVLSACASLSGLSQGGEMHVIYIDKRG IETEGFLATALVDMYSKCGIKDKALEYFRATSKRDVSTWNSTISLSVHGIGKDALDI FSEMYVEGFKPNGITFIGVLSACNHVGMLDQARKLEFMMSVYRVEPTIEHYGCMVDL LGRMGKIEEAEELVNEIPADEASILLESLLGACKRFQGLEQAERIANRLLLENLRRSS GYQMNSLYASDGRWEKVIDGRNMRNRAERVNRSLDVA"					
complement(6832..7011)					
/gene="AT4g18850"					
/number=1					
complement(6832..7011)					
/gene="AT4g18850"					
gene	6832..7011				
	/gene="AT4g18850"				
CDS	complement(6832..7011)				
	/gene="AT4g18850"				
	/codon_start=1				
	/product="hypothetical protein"				
	/protein_id="CAAL6742.1"				
	/db_xref="GI:2832613"				
	/db_xref="SPTREMBL:O49400"				
	/translation="MVDLDLQQIROLLENIINHLEANELLPPQPNMYPHDPLEDLGDS PASSDVTEGNSDED"				
gene	9456..10357				
	/gene="AT4g18860"				
CDS	join(9456..9560,10169..10357)				
	/gene="AT4g18860"				
	/codon_start=1				
	/product="hypothetical protein"				
	/protein_id="CAAL6743.2"				
	/db_xref="GI:5922639"				
	/db_xref="SPTREMBL:O49401"				
	/translation="MPSVQVRMDNKTAAELDAIENLLKCCIDVMTITTKHNLALQRP SFITNYPKALYFDSSSAVSOLDGFTSEGTASTRVLGKSASWIFFNFVTEK"				
exon	9456..9560				
	/gene="AT4g18860"				
	/number=1				
intron	9561..10168				
	/gene="AT4g18860"				
	/number=1				
exon	10169..10357				
	/gene="AT4g18860"				
	/number=2				
gene	10938..11996				
	/gene="AT4g18870"				
exon	10938..11150				
	/gene="AT4g18870"				
	/number=1				
CDS	join(10938..11150,11238..11639,11736..11996)				
	/gene="AT4g18870"				
	/note="similarity to heat shock transcription factor HSFL1, Arabidopsis thaliana, PIR2:S52641				
	contains Prenyl group binding site (CAAX box) AA289-292"				
	/codon_start=1				
	/product="heat shock transcription factor-like protein"				
	/protein_id="CAAL6744.1"				
	/db_xref="GI:2832615"				
	/db_xref="SPTREMBL:O49402"				
	/translation="MSKNEGSLTSVSVEITTYDMVDDLSDLSTISWSQSGKSPFIIN PEEFYNNLQRCFCORINTFFSFLSGCFRKIDSGAKFEANDNFVRCGRHLINNIISD VIQEYQIQDQWELFKAEKFLARLVKAVQDLQPLPHNSYFTSKLPFTTILYEMVDPPSS DALIISWSQSGKFLIINWQPEFCKDHLRRLPNTLUIHFFYFKLKIFGPKKINPKWFEA NDNFVRCGRHLVEIILISNDKKNDQLRKQDAREKMAEAGELFKIQIEMSDMRKKMK				

VELAKLVSSDDHALSVSRQFGLMDVSAKSNLIRSLKRVGSLRNLPEPTITNQENTN
KSISSGDADGDIYRKDEIESYSSRNSSEELTESSLSLTVSRVPRKPPKRSISL
GDSSTENRADPPQKSIPIPPPPPPPLQLQPPPPSVSKAPPPPPPPPKSLSIASA
KVRVPVEVFEYHSLMRDSTNSRDRSTGGNAAAEAILANSNARDMIGEIEENRSVYL
LAITDVTETGDFIRELIKEVGNAAFSDIEDVYFVKWLDLDSLYLFKVCXFFVVSXK
VDERAVLKHFENPEQKADALREAPFCYFDLKKLISBRASRFREDPROSSSALKKMOAL
FEKLEHGVSLSBMRRESATKPKFSQIPVDMMLETGTITSOIKLASVKLAMKRVSA
ELEALEGGPEEBELIVQGVRAFRVHQFAGGDAETMKAFEBELRDKARSCHVQCQSQ
THQKLCFRSRIPC
9577. .9906
/gene="AT4g18560"
/number=1
9612. .11059
/gene="AT4g18570"
complement(join(9612. .9715,10483. .10747,10973. .11059))
/gene="AT4g18570"
complement(join(9612. .9715,10483. .10747,10973. .11059))
/gene="AT4g18570"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB78859.1"
/db_xref="GI:7268651"
/translation="MSENAAFTSLIKNLMKSPCVSTVFICNQVLLPRRFRHRWML
FSSLWSLFASTSDKTOPLPVLPFLSPTOCLNSSAAVVERAEELAKQTEVVVVVEEAV
AEVEEEAESISAVAREPEGLSGGGGGRKGEGHSLFLVEGDL"
complement(9612. .9715)
/gene="AT4g18570"
/number=1
complement(9716. .10482)
/gene="AT4g18570"
/number=1
9907. .9957
/gene="AT4g18570"
/number=1
9958. .10809
/gene="AT4g18570"
/number=2
complement(10483. .10747)
/gene="AT4g18570"
/number=2
complement(10748. .10972)
/gene="AT4g18570"
/number=2
10810. .10979
/gene="AT4g18570"
/number=2
complement(10973. .11059)
/gene="AT4g18570"
/number=3
10980. .11111
/gene="AT4g18560"
/number=3
11112. .11159
/gene="AT4g18560"
/number=3
11160. .11377
/gene="AT4g18560"
/number=4
11378. .11462
/gene="AT4g18560"
/number=4
11463. .11574
/gene="AT4g18560"
/number=5
11575. .12309
/gene="AT4g18560"
/number=5
12310. .12447
/gene="AT4g18560"
/number=6
12448. .12540
/gene="AT4g18560"
/number=6

12541. .13672
/gene="AT4g18560"
/number=7
complement(join(12991. .13077,13176. .13264,13382. .13507,
13596. .13662))
/gene="AT4g18580"
12991. .13662
/gene="AT4g18580"
complement(join(12991. .13077,13176. .13264,13382. .13507,
13596. .13662))
/gene="AT4g18580"
/note="contains EST gb:AI998801.1, AI138148"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB78860.1"
/db_xref="GI:7268652"
/translation="MATLAPVHAVSVLHSHRRKGGCNGNFVSGVPLPYLIPRRKPI
PVDFDYNNSSDESGEGDKGSHSDEYVVVDKAEAIMPNVKTGKPGAIQMELSKDAV
GDDESKAKDAGYSDEWFFV"
complement(12991. .13077)
/gene="AT4g18580"
/number=1
complement(13078. .13175)
/gene="AT4g18580"
/number=1
complement(13176. .13264)
/gene="AT4g18580"
/number=2
complement(13265. .13381)
/gene="AT4g18580"
/number=2
complement(13382. .13507)
/gene="AT4g18580"
/number=3
complement(13508. .13595)
/gene="AT4g18580"
/number=3
complement(13596. .13662)
/gene="AT4g18580"
/number=4
14662. .20681
/gene="AT4g18590"
join(14662. .14917,15014. .15052,16278. .16535,16652. .167,
17087. .17423,17520. .17854,18312. .18366,19447. .19527,
19615. .20681)
/gene="AT4g18590"
/note="overlap to next genmodel
similarity to 18.3K protein precursor, pollen, sea mays,
PIR2:J01107
Contains Pollen proteins Ole e I family signature
AA309-318
contains EST gb:AI999674.1, AA721906"
/codon_start=1
/product="pollen-specific protein-like"
/protein_id="CAB78861.1"
/db_xref="GI:7268653"
/translation="MDTSPAAFVNGALLRRYIQCKVRVAVIQVIRSDVSGSVIGKSTDD
QQIIVKGSPPQPLTITLLEVGIAETDNTIRAEVWTFEGSDFDQVNTNCLCKLANDRRK
AMTDLQMEVEVDNTSSQLSPKQVNYRCKKRRIVAIEENIVPHEPKGEECFAWK
KRSGNSEQVQCSSIFVEPHMKMQTIHDGMVEEKLLCFGCNGRILGYFNWAGMQSCGAW
LKITDGFNFMILYKTELCSGSELVYLFNSINIKENDYKREEVSKRRPESFNILKQMAS
KAIFFFVSACLSSLAGVAIADADDLDFRQIQGSVYCDTCRVQVFTRLSKLEGAKY
KLECRSTNGTVTLTKREAVTDKTSYRMEVTDGHEEVEVCBLVLPSPDSCSDSVSKEA
YLNAAKISLTANDGIVNTHETIRVNPGLFVQVTPSAECPAARFELGIVPDGLLFGWFG
GOARNISETDGRSVTTHGAATCISGRFKEVSESTSAKARYAAGSGRIFISLGTGS
SGKHLCEPVSVUKLGTPEPIFSSKTASVEDRSVIVSLGVLGGLVSVVLELEASE
FOIFSSKTASFERSVAKGFSKVLKCRPVSAEPGASEPGIISPKTASFKEISVATG

Query Match 5.9%; Score 160.2; DB 8; Length 199075;
Best Local Similarity 67.6%; Pred. No. 5.5e-17;
Matches 225; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
Qy 1516 gg-gatgtgatgtcatattgtggaggccaccatgccaaaggtatcagtggttataatcgg 1575

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model.

Run on: July 15, 2002, 09:12:48 ; Search time 425.1 Seconds
(without alignments)
11050.289 Million cell updates/sec

Title: US-09-914-001-1

Perfect score: 2736

Sequence: 1 atggcgccagctcccgctc.....caggggagtagtgtagcag 2736

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2736	100.0	2736	21	Zea mays zmet2a me
2	2736	100.0	2736	21	Zea mays zmet2a me
3	616.4	22.5	1181	21	Zea mays zmet2b me
4	466.8	17.1	12810	21	Zmet2a methyltrans
5	154.4	5.6	520	22	Arabidopsis thalia
6	83.2	3.0	795	19	FUGA insert stabil
7	81.2	3.0	799	19	Nucleotide sequenc
8	78.6	2.9	1926	21	Epstein Barr virus
9	78.6	2.9	1926	22	EBV tethering prot

c 10	78.6	2.9	2580	21	AA75454
c 11	78.6	2.9	5452	20	AA90923
c 12	78.6	2.9	8705	20	AA23778
c 13	78.6	2.9	9600	19	AAV21683
c 14	78.6	2.9	10380	20	AAZ2248
c 15	78.6	2.9	10596	14	AAQ51731
c 16	78.6	2.9	10596	17	AA40348
c 17	78.6	2.9	10596	20	AA15650
c 18	78.6	2.9	16080	21	AA59553
c 19	76.2	2.8	51259	18	AA83007
c 20	75.4	2.8	16442	18	AA83006
c 21	74.6	2.7	1925	20	AA90924
c 22	73.6	2.7	33923	22	AA67071
c 23	71	2.6	1527	23	AA68546
c 24	71	2.6	1620	23	AA68547
c 25	71	2.6	2178	23	AA86106
c 26	70.8	2.6	693	23	AA574240
c 27	70.8	2.6	693	23	AA590715
c 28	70.4	2.6	3489	21	AA30290
c 29	70.4	2.6	3489	22	AA82901
c 30	70.4	2.6	32207	20	AAV73805
c 31	70.4	2.6	137507	19	AAV19941
c 32	69.6	2.5	267	22	AAK19599
c 33	69.6	2.5	267	22	AAK45604
c 34	69.6	2.5	474	22	AAK06682
c 35	69.6	2.5	474	22	AAK32380
c 36	69.4	2.5	1548	19	AAV33277
c 37	69.4	2.5	1581	19	AAV33278
c 38	69.2	2.5	154746	24	AA25519
c 39	69.2	2.5	154746	24	AA25519
c 40	69	2.5	125157	22	AAH74202
c 41	68.4	2.5	6313	13	AB199537
c 42	68	2.5	1080	24	AB199537
c 43	68	2.5	117213	19	AAV62176
c 44	67.2	2.5	124884	22	AAH74201
c 45	66.8	2.4	2277	19	AAV13836

ALIGNMENTS

RESULT 1	
AAA95233	
ID	AAA95233 standard; cDNA; 2736 BP.
XX	AC
XX	AAA95233;
XX	21-FEB-2001 (first entry)
DT	
DE	Zea mays zmet2a methyltransferase coding sequence #1.
XX	
XX	Maize; corn; zmet2a; DNA methyltransferase; cytosine-5-methyltransferase;
KW	DNA methylation; marker-aided selection; chromosome segregation;
KW	DNA fingerprinting; gene therapy; trinucleotide repeat; ss.
XX	
XX	Zea mays.
OS	
XX	
XX	Location/Qualifiers
FT	1..2736
FT	/*tag= a
FT	/product= "zmet2a #1"
FT	/partial
XX	
XX	WO200053732-A2.
PD	
PD	14-SEP-2000.
XX	
XX	10-MAR-2000; 2000WO-US06456.
XX	
XX	11-MAR-1999; 99US-0123888.
PR	09-DEC-1999; 99US-0169858.
XX	
PA	(WISC) WISCONSIN ALUMNI RES FOUND.

PA (PION-) PIONEER HI-BRED INT INC.
PA (MINU) UNIV MINNESOTA.

PI Kaeppler SM, Springer NM, Muszynski MG, Papa CM;

DR WPI; 2000-587430/55.

Novel zmet2a and zmet2b methyltransferase nucleic acid sequences of Zea mays used for stabilizing transgene expression, altering yield or biochemical qualities, and silencing targeted genes in transgenic plants

XX
PS
Claim 2: Fig 1A: 103pp: English.

The present sequence is one version of the coding sequence for the maize cytosine-5-methyltransferase *zmaMET1*. The protein encoded by this sequence is involved in the methylation of DNA, which is necessary for cell growth and development in plants and is used in the selection of transgenic plants with altered levels of DNA methylation in the production of transgenic plants with altered levels of DNA methylation in gene silencing in marker-aided selection, to follow chromosome segregation, and to DNA fingerprint class II DNA methyltransferases in other plants. In addition, they can be used in gene therapy to treat human diseases caused by trinucleotide repeat amplification.

Sequence 2736 BP: 684 A; 628 C; 802 G; 622 T; 0 other;

```
Query Match      100.0%; Score 2736; DB 21; Length 2736;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 atggcgcgagctccccggtcacccgcgcgcctacacgcgtctctctgggcggaagcgcgcc 60

Db 1 atggcgcgagctcccggtcacccgcgcgcctacacgcgtctctgggagcgcc 60

QY 61 gccaggccgaggagatccaccagaacaaggaggaggagggtcgcggcggtcc 120

Db 61 gccaggccgaggaatcaccagaacaaggaggaggagggtcgcgggcgctcc 120

Qy 121 tccgccaagcgcagccgcgaaggcgcatcttccgggaagaagcccgaagtcccccccaag 180

Db 121 tccgcaagcgagccgcaaggcgcatcttcgggaagagcccaagtccccccaag 180

Qy 181 caggccaagccggggagggaagaaggggatgccgagatgaaggagcccgaggagac 240

Db 181 caggccaagccggggagggaagaaaggggatgccgagatgaaggagcccggtggaggac 240

QY 241 gacgtgtgcgcgaggagcccgcagcagaggaggttgcccatgggcgagggagggccgag 300

Dd 241 gacgtgtgcgccgagggagcccgacgagggaggttggccatggcgagggagggccgag 300

Qy	601	gcagattacaactggccgcattactactgaatttttttggagggaactgacccagtgctcaatttt	660
Db	601	gcagattacaactggccgcattactactgaatttttttggagggaactgacccagtgctcaatttt	660
Qy	661	acttgcgcgttggtttctccagcagagagacacggttatcaatttttggctgcataagt	720
Db	661	acttgcgcgttggtttctccagcagagagacacggttatcaatttttggctgcataagt	720
Qy	721	gtggatggccacaagaatgacccctagaagctgtttttctctcgaggaanaagacacaaat	780
Db	721	gtggatggccacaagaatgacccctagaagctgtttttctctcgaggaanaagacacaaat	780
Qy	781	gtgcttgattgatttatctccaaggtccaagaatccaatgttgatccaaatatgatcca	840
Db	781	gtgcttgattgatttatctccaaggtccaagaatccaatgttgatccaaatatgatcca	840
Qy	841	aaagccaaaggtcagctgtagagaggttgcagacctatactatatgacatgtcttactctgtt	900
Db	841	aaagccaaaggtcagctgtagagaggttgcagacctatactatatgacatgtcttactctgtt	900
Qy	901	gcattctcacatttgcataatctctgctctgaaaaagggcagtcaggcagtgataccgct	960
Db	901	gcattctcacatttgcataatctctgctctgaaaaagggcagtcaggcagtgataccgct	960
Qy	961	tggggtattttctctgtagtgatctggagagcttcactagtagtgcacaaagaggaca	1020
Db	961	tggggtattttctctgtagtgatctggagagcttcactagtagtgcacaaagaggaca	1020
Qy	1021	gcaacccctcttgactgtattcttgagctgtgggggaatgtctacttggtcttcttgaggt	1080
Db	1021	gcaacccctcttgactgtattcttgagctgtgggggaatgtctacttggtcttcttgaggt	1080
Qy	1081	gcagctctttctggcttgaacctggaactcgatgggctgttgatttcaacagttttgcg	1140
Db	1081	gcagctctttctggcttgaacctggaactcgatgggctgttgatttcaacagttttgcg	1140
Qy	1141	tgccaaagttttaaataataatcatccacagactgaggtgcgaaaatgagaaagccgatgag	1200
Db	1141	tgccaaagttttaaataataatcatccacagactgaggtgcgaaaatgagaaagccgatgag	1200
Qy	1201	ttctcttgccctcctcaagaagaatgggcagttctatgcacaaaataatgtccaagatgtgag	1260
Db	1201	ttctcttgccctcctcaagaagaatgggcagttctatgcacaaaataatgtccaagatgtgag	1260
Qy	1261	tcaaattagcaagctcagagatcaagcgatcaagcagacagccctcttgacaagacacaa	1320
Db	1261	tcaaattagcaagctcagagatcaagcgatcaagcagacagccctcttgacaagacacaa	1320
Qy	1321	tttgtgtgagagaacttgctggggataatgtatgtgtggcagtgacaggggaaaaatggcatc	1380
Db	1321	tttgtgtgagagaacttgctggggataatgtatgtgtggcagtgacaggggaaaaatggcatc	1380
Qy	1381	tatttttaagggtccagtggaagagatacagggccctgaggaggaacatatgggaaccgattgat	1440
Db	1381	tatttttaagggtccagtggaagagatacagggccctgaggaggaacatatgggaaccgattgat	1440
Qy	1441	aacttgagtgactgcccgcagaaaaattagaaaaattgtacaaagagggcacaaaagaag	1500
Db	1441	aacttgagtgactgcccgcagaaaaattagaaaaattgtacaaagagggcacaaaagaag	1500
Qy	1501	attctccactgactgattgtattgtattgtattgtggaggccaccatgtccaagtatcc	1560
Db	1501	attctccactgactgattgtattgtattgtattgtggaggccaccatgtccaagtatcc	1560
Qy	1561	agtgggtttaatctgttacagaaaccgtgatgagcactcaaatagatgagaaaaacaacaa	1620
Db	1561	agtgggtttaatctgttacagaaaccgtgatgagcactcaaatagatgagaaaaacaacaa	1620
Qy	1621	atggtgactttcaatggatatgttggcgctacttgagcccaagtagttcttccatggaaaaat	1680
Db	1621	atggtgactttcaatggatatgttggcgctacttgagcccaagtagttcttccatggaaaaat	1680
Qy	1681	qtggtgacatactcaaatttgcggatgttaacctagagaaaaatagtctttgagctgcctt	1740

```

|||||
Db 1681 gtgttgacatactcaaatctcgatggttaactagaaaaatatgctttagtgcctt 1740
|||
Qy 1741 gttgctatgaagtaccagcggtctggaatgatggtgctggttgcctatggtctgcca 1800
|||||
Db 1741 gttgctatgaagtaccagcggtctggaatgatggtgctggttgcctatggtctgcca 1800
|||||
Qy 1801 cagttcaagatgcgtgttctctcgtggtgctctcttcttccatggtgctcccaagtat 1860
|||||
Db 1801 cagttcaagatgcgtgttctctcgtggtgctctcttcttccatggtgctcccaagtat 1860
|||||
Qy 1861 cctctgccacactatgatgttagtcgtgagagaccctaaatgccttttcgaatgt 1920
|||||
Db 1861 cctctgccacactatgatgttagtcgtgagagaccctaaatgccttttcgaatgt 1920
|||||
Qy 1921 atggttgcatatgacgagacacaaaaacatccctgaaaaaacctctctcttctgagcat 1980
|||||
Db 1921 atggttgcatatgacgagacacaaaaacatccctgaaaaaacctctctcttctgagcat 1980
|||||
Qy 1981 gcaatttcagatttaccaaaaggttcaaaatcacccagcctcaacgatgtgatgagtatggt 2040
|||||
Db 1981 gcaatttcagatttaccaaaaggttcaaaatcacccagcctcaacgatgtgatgagtatggt 2040
|||||
Qy 2041 ggttcccccaagaccgaattccagcgtacattcgcactcagtcgtaaaagacatgttgat 2100
|||||
Db 2041 ggttcccccaagaccgaattccagcgtacattcgcactcagtcgtaaaagacatgttgat 2100
|||||
Qy 2101 tggctctctggtgaggggctggtccagatgaaggcaagctcttggatcacccagccttta 2160
|||||
Db 2101 tggctctctggtgaggggctggtccagatgaaggcaagctcttggatcacccagccttta 2160
|||||
Qy 2161 cggcttaacaacgatattatgagcgggttcaacagattctctcgaaggagggccaac 2220
|||||
Db 2161 cggcttaacaacgatattatgagcgggttcaacagattctctcgaaggagggccaac 2220
|||||
Qy 2221 ttccgcgacctaaaggcgtgaggttgagcaacaataattgtgagtgagatccagaa 2280
|||||
Db 2221 ttccgcgacctaaaggcgtgaggttgagcaacaataattgtgagtgagatccagaa 2280
|||||
Qy 2281 atcgagcgtgtgaaccttctatctggaacacactggttctgactgactgcaatgtcatc 2340
|||||
Db 2281 atcgagcgtgtgaaccttctatctggaacacactggttctgactgactgcaatgtcatc 2340
|||||
Qy 2341 atcaaggcacaatcactcaagcgttggcgctgtggtggagacgagacagattcctaca 2400
|||||
Db 2341 atcaaggcacaatcactcaagcgttggcgctgtggtggagacgagacagattcctaca 2400
|||||
Qy 2401 gttgtaaccagacagagcctcacaaccaggttaataattccatccgactcaagcagggtc 2460
|||||
Db 2401 gttgtaaccagacagagcctcacaaccaggttaataattccatccgactcaagcagggtc 2460
|||||
Qy 2461 ctcaactatccggggaacgcaagggttaacagggtctcccgattactaccgattgtttggc 2520
|||||
Db 2461 ctcaactatccggggaacgcaagggttaacagggtctcccgattactaccgattgtttggc 2520
|||||
Qy 2521 ccgatacagagaagtagtattcaagtcggggaacgagtgctgctccctgttcccgaggca 2580
|||||
Db 2521 ccgatacagagaagtagtattcaagtcggggaacgagtgctgctccctgttcccgaggca 2580
|||||
Qy 2581 ctgggctactgctctggggaacgctactcgtggtgaattctgagggagtgacctctgtac 2640
|||||
Db 2581 ctgggctactgctctggggaacgctactcgtggtgaattctgagggagtgacctctgtac 2640
|||||
Qy 2641 cagctgcctcaagtctcaactctgttgaggagcagcactgcggggcagcgaggcctct 2700
|||||
Db 2641 cagctgcctcaagtctcaactctgttgaggagcagcactgcggggcagcgaggcctct 2700
|||||
Qy 2701 cctgttggcacccctgcaggggaggttagttgagcag 2736
|||||
Db 2701 cctgttggcacccctgcaggggaggttagttgagcag 2736
|||||
```

RESULT 2

```

AAA95234
ID AAA95234 standard; cDNA; 2796 BP.
XX
AC AAA95234;
XX
DT 21-FEB-2001 (first entry)
XX
DE Zea mays zmet2a methyltransferase coding sequence #2.
XX
KW Maize; corn; zmet2a; DNA methyltransferase; cytosine-5-methyltransferase;
KW DNA methylation; marker-aided selection; chromosome segregation;
KW DNA fingerprinting; gene therapy; trinucleotide repeat; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 1..2796
FT /tag= a
FT /product= "zmet2a #2"
FT /transl_except= (pos:2767..2769,aa:Xaa)
FT /note= "this sequence contains an in-frame stop codon
FT represented by Xaa in the protein sequence"
FT /partial
XX
PN WO200053732-A2.
XX
PD 14-SEP-2000.
XX
PF 10-MAR-2000; 2000WO-US06456.
XX
PR 11-MAR-1999; 99US-0123888.
PR 09-DEC-1999; 99US-0169858.
XX
PA (WISC ) WISCONSIN ALUMNI RES FOUND.
PA (PION-) PIONEER HI-BRED INT INC.
PA (MINU ) UNIV MINNESOTA.
XX
PI Kaeppler SM, Springer NM, Muszynski MG, Papa CM;
XX
DR WPI: 2000-587430/55.
DR P-PSDB: AAB26243.
XX
PT Novel zmet2a and zmet2b methyltransferase nucleic acid sequences of Zea
PT mays used for stabilizing transgene expression, altering yield or
PT biochemical qualities, and silencing targeted genes in transgenic
PT plants -
XX
PS Claim 4; Fig 1B; 103pp; English.
XX
CC The present sequence is one version of the coding sequence for the maize
CC cytosine-5-methyltransferase zmet2a. The protein encoded by this sequence
CC is involved in the methylation of DNA, which is necessary for normal
CC development in plants and animals. The sequence, protein, antibodies and
CC antisense sequences can be used in the production of transgenic plants
CC with altered levels of DNA methylation, in gene silencing, in
CC marker-aided selection, to follow chromosome segregation, and to DNA
CC fingerprint class II DNA methyltransferases in other plants. In
CC addition, they can be used in gene therapy to treat human diseases
CC caused by trinucleotide repeat amplification.
XX
SQ Sequence 2796 BP; 701 A; 645 C; 820 G; 630 T; 0 other;
```

```

Query Match 100.0%; Score 2736; DB 21; Length 2796;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atggcgccgagctcccgctaccccgccgctacacgctctctggcggaagcgcc 60
|||
Db 31 atggcgccgagctcccgctaccccgccgctacacgctctctggcggaagcgcc 90
|||
Qy 61 gccaaagccagagagatccaccagacaagagagagagaggtcgcgcgcgtcc 120
|||
```

D	b	91	gcaagcgagagatccaccagaacaggagaggagggaggtcgcggcgctcc	150
Q	y	121	tcgcgaagcgagcgcaagcgcatcttcgggaagaagcccaagtgcgcccccaag	180
D	b	151	tcgcgaagcgagcgcaagcgcatcttcgggaagaagcccaagtgcgcccccaag	210
Q	y	181	caggccaagcgaggaagaagaaggggatgcgagatgaagagcccgtagagac	240
D	b	211	caggccaagcgaggaagaagaaggggatgcgagatgaagagcccgtagagac	270
Q	y	241	gacgtgtcgcccagagagcccgacgagaggagttggccatggcgagggagccgag	300
D	b	271	gacgtgtcgcccagagagcccgacgagaggagttggccatggcgagggagccgag	330
Q	y	301	gagcagggcatcgagagagggtggttcggttcgcgcggggtcaccccggaagaagag	360
D	b	331	gagcagggcatcgagagagggtggttcggttcgcgcggggtcaccccggaagaagag	390
Q	y	361	gtgggagaagaacgcgcgcgcgcgtgcgaccccgagcccgaggttcacgcgcgc	420
D	b	391	gtgggagaagaacgcgcgcgcgcgtgcgaccccgagcccgaggttcacgcgcgc	450
Q	y	421	ccgtgtgcgcgagagcgcgagcaactggccaaagcgctacggccgacactgcc	480
D	b	451	ccgtgtgcgcgagagcgcgagcaactggccaaagcgctacggccgacactgcc	510
Q	y	481	gcaagaacccgatgagaggaagagctcaagggcagatcactaccgagcgctaag	540
D	b	511	gcaagaacccgatgagaggaagagctcaagggcagatcactaccgagcgctaag	570
Q	y	541	gtggcaaacgctcgtctactgcctcggggatgacgtctctatgtcaaggctggaaaaacgag	600
D	b	571	gtggcaaacgctcgtctactgcctcggggatgacgtctctatgtcaaggctggaaaaacgag	630
Q	y	601	gcagattacattggcgcattactgaatttttgaggggactgacagtgctactattt	660
D	b	631	gcagattacattggcgcattactgaatttttgaggggactgacagtgctactattt	690
Q	y	661	acttgccgttgatttccgagcagagcaaggttatcaatttttggtgtccataagt	720
D	b	691	acttgccgttgatttccgagcagagcaaggttatcaatttttggtgtccataagt	750
Q	y	721	gtggatggccacaagcatgacctagacgtgtttttcttcttgagaaaagacgacaat	780
D	b	751	gtggatggccacaagcatgacctagacgtgtttttcttcttgagaaaagacgacaat	810
Q	y	781	gtgcttgattgcatatctccaaggtcaagatagtcacatgttgatccaatatggatcca	840
D	b	811	gtgcttgattgcatatctccaaggtcaagatagtcacatgttgatccaatatggatcca	870
Q	y	841	aaagccaaggtcagctgatagagagttgcgacctatactatgacatgttactctgtt	900
D	b	871	aaagccaaggtcagctgatagagagttgcgacctatactatgacatgttactctgtt	930
Q	y	901	gcattctacatttgcataatctcgtctgaaaaatggcgagtcaggcagtgataccgct	960
D	b	931	gcattctacatttgcataatctcgtctgaaaaatggcgagtcaggcagtgataccgct	990
Q	y	961	tcgggtatttctctgatgtgagatcttgagacgtcatctagtatgccaagagaca	1020
D	b	991	tcgggtatttctctgatgtgagatcttgagacgtcatctagtatgccaagagaca	1050
Q	y	1021	gcaacctctctgatctgtattctgctgtggggcgatgctactggtcttcttggtgt	1080
D	b	1051	gcaacctctctgatctgtattctgctgtggggcgatgctactggtcttcttggtgt	1110
Q	y	1081	gcagctcttctcggcttgaaacttgaaactcgatggctgtgttgatttcaacagttttgcg	1140
D	b	1111	gcagctcttctcggcttgaaacttgaaactcgatggctgtgttgatttcaacagttttgcg	1170
Q	y	1141	tgccaagtttataataataatcatccacagactgggtgcgaatatgagaagccatag	1200
D	b	1171	tgccaagtttataataataatcatccacagactgggtgcgaatatgagaagccatag	1230
Q	y	1201	ttcttgccctccttaagaaatgggcagttctatgcaaaaaatgttccaagatgtggat	1260
D	b	1231	ttcttgccctccttaagaaatgggcagttctatgcaaaaaatgttccaagatgtggat	1290
Q	y	1261	tcaaaattagcaagctcagagatcaagcgagatgaagcagccctcttgacaagagcaga	1320
D	b	1291	tcaaaattagcaagctcagagatcaagcgagatgaagcagccctcttgacaagagcaga	1350
Q	y	1321	tttggttagaagaacttgtcgggatatgttatgtgtgcagtgacagggaaaaatggcatc	1380
D	b	1351	tttggttagaagaacttgtcgggatatgttatgtgtgcagtgacagggaaaaatggcatc	1410
Q	y	1381	tattttaaagttccagtggaagagatcggccctgagggagatatacatgggaacgattgat	1440
D	b	1411	tattttaaagttccagtggaagagatcggccctgagggagatatacatgggaacgattgat	1470
Q	y	1441	aacttgagtgactgcgcgcagaaaaattagagaatttgtacagaagggcacaagaagaag	1500
D	b	1471	aacttgagtgactgcgcgcagaaaaattagagaatttgtacagaagggcacaagaagaag	1530
Q	y	1501	attctccactgcctgggtgatgtgtgatgttattgttgaggcccaaccatgccaaagtatc	1560
D	b	1531	attctccactgcctgggtgatgtgtgatgttattgttgaggcccaaccatgccaaagtatc	1590
Q	y	1561	agtgggttttaactgcgtacagaaaaccgtgatgagccactcaaaagatgagaaaaacaacaa	1620
D	b	1591	agtgggttttaactgcgtacagaaaaccgtgatgagccactcaaaagatgagaaaaacaacaa	1650
Q	y	1621	atggtgaacttcatgatatattgtggtgacttgagagcccaagatgtttctcatggaaaaat	1680
D	b	1651	atggtgaacttcatgatatattgtggtgacttgagagcccaagatgtttctcatggaaaaat	1710
Q	y	1681	gtggtggacatactcaaaatttgcgagatggtttacactagagaaaaatgctttgagctgcctt	1740
D	b	1711	gtggtggacatactcaaaatttgcgagatggtttacactagagaaaaatgctttgagctgcctt	1770
Q	y	1741	gttgctatgaagtccaagcgcgcttggaatgatgtggctgtgtctgtctgtctgccca	1800
D	b	1771	gttgctatgaagtccaagcgcgcttggaatgatgtggctgtgtctgtctgtctgccca	1830
Q	y	1801	cagttcagagatcggtgttctctctgggtgtctctttcttccatggctgcctccatagtat	1860
D	b	1831	cagttcagagatcggtgttctctctgggtgtctctttcttccatggctgcctccatagtat	1890
Q	y	1861	cctctgcccacactatgatgtgtgtagtacgtggagagccctaatgctcttttcgcaatgt	1920
D	b	1891	cctctgcccacactatgatgtgtgtagtacgtggagagccctaatgctcttttcgcaatgt	1950
Q	y	1921	atggttgcatatgcgagacacaaaaaacctacccctgaaaaagccttgcttcttgccgat	1980
D	b	1951	atggttgcatatgcgagacacaaaaaacctacccctgaaaaagccttgcttcttgccgat	2010
Q	y	1981	gcaatttcgaatttaccagaaggttcaaaatccaccagctcaacgtatgataggatgat	2040
D	b	2011	gcaatttcgaatttaccagaaggttcaaaatccaccagctcaacgtatgataggatgat	2070
Q	y	2041	ggttcccccagaacccgaattccagcgctacattcgactcagtcgtataagacatgttggat	2100
D	b	2071	ggttcccccagaacccgaattccagcgctacattcgactcagtcgtataagacatgttggat	2130
Q	y	2101	tggtcctctcgggtgaaggggctggtccagatgaaggcaagccttggtatccagcccttca	2160
D	b	2131	tggtcctctcgggtgaaggggctggtccagatgaaggcaagccttggtatccagcccttca	2190
Q	y	2161	cggcttacaacagatgattatgacggggttcaacagattcctgtcaagaaggagggcacaac	2220
D	b	2191	cggcttacaacagatgattatgacggggttcaacagattcctgtcaagaaggagggcacaac	2250
Q	y	2221	ttccgcgacctaaaagcggtgaggggttgagcgaacaaatatgtttgagtggtgtccagaa	2280
D	b	2251	ttccgcgacctaaaagcggtgaggggttgagcgaacaaatatgtttgagtggtgtccagaa	2310


```

|||||  || || || ||| ||||| ||| || ||||| ||||| |||||
Db 705 acctctgt---ggggcgaacgggttcaggggaatgcgcttctgttgccactctgcg 761
QY 2719 ggggaggttagtgagcag 2736
|||||  |||||
Db 762 ggggaggttagtcagcag 779

RESULT 4
AAA95272
ID AAA95272 standard; DNA; 12810 BP.
XX
AC AAA95272;
XX
DT 21-FEB-2001 (first entry)
XX
DE Zmet2a methyltransferase gene-retrotransposon SPRITE-1 fusion construct.
XX
KW Maize; corn; zmet2a; DNA methyltransferase; cytosine-5-methyltransferase;
KW DNA methylation; marker-aided selection; chromosome segregation;
KW DNA fingerprinting; gene therapy; trinucleotide repeat; fusion construct;
KW SPRITE-1 retrotransposon; ss.
XX
OS Zea mays.
OS Synthetic.
XX
FH Key
FH primer_bind
FT 747..768 Location/Qualifiers
FT /*tag= a
FT /bound_moiety= "primer 17F"
FT 792..809
FT /*tag= b
FT /bound_moiety= "primer 17R"
FT 1171..1196
FT /*tag= c
FT /bound_moiety= "primer 9eF"
FT complement (1290..1311)
FT /*tag= d
FT /bound_moiety= "primer 9eR"
FT 1762..1783
FT /*tag= e
FT /bound_moiety= "primer 161F"
FT 2571..2595
FT /*tag= f
FT /bound_moiety= "primer 151F"
FT 2899..2919
FT /*tag= g
FT /bound_moiety= "primer 14eF"
FT complement (2899..2919)
FT /*tag= h
FT /bound_moiety= "primer 15eR"
FT complement (4467..4494)
FT /*tag= i
FT /bound_moiety= "primer 181R"
FT complement (5043..5068)
FT /*tag= j
FT /bound_moiety= "primer 161R"
FT complement (5404..5428)
FT /*tag= k
FT /bound_moiety= "primer 141R"
FT 5648..5674
FT /*tag= l
FT /bound_moiety= "primer 111F"
FT 5709..5734
FT /*tag= m
FT /bound_moiety= "primer 121F"
FT complement (6867..6890)
FT /*tag= n
FT /bound_moiety= "primer 131R"
FT 7081..7102
FT /*tag= o
FT /bound_moiety= "primer 131F"
FT complement (7334..7354)
FT

/*tag= p
/bound_moiety= "primer 541R"
complement (7381..7408)
/*tag= q
/bound_moiety= "primer 121R"
complement (7757..7779)
/*tag= r
/bound_moiety= "primer 56RN"
complement (8037..8053)
/*tag= s
/bound_moiety= "primer 111R"
8196..8219
/*tag= t
/bound_moiety= "primer 7F"
8281..8304
/*tag= u
/bound_moiety= "primer 8eF"
complement (8284..8312)
/*tag= v
/bound_moiety= "primer 8eR"
complement (8370..8391)
/*tag= w
/bound_moiety= "primer 5RN"
8813..8836
/*tag= x
/bound_moiety= "primer Seq2FN"
9238..9259
/*tag= y
/bound_moiety= "primer 531F"
complement (10062..10085)
/*tag= z
/bound_moiety= "primer Seq2RN"
complement (10381..10404)
/*tag= aa
/bound_moiety= "primer Seq1RN"
complement (10440..10460)
/*tag= ab
/bound_moiety= "primer 53eR"
10654..10675
/*tag= ac
/bound_moiety= "primer 3F"
complement (10664..10684)
/*tag= ad
/bound_moiety= "primer 3R"
10860..10885
/*tag= ae
/bound_moiety= "primer 1F"
10897..10923
/*tag= af
/bound_moiety= "primer Race2A"
complement (10897..10923)
/*tag= ag
/bound_moiety= "primer Race2B"
11019..11057
/*tag= ah
/bound_moiety= "primers Race1A and 4F"
complement (11019..11059)
/*tag= ai
/bound_moiety= "primers Race1B and RaceRT"
11756..11775
/*tag= aj
/bound_moiety= "primer 5F"
complement (12731..12758)
/*tag= ak
/bound_moiety= "primer 1R"

WO200053732-A2.
14-SEP-2000.
10-MAR-2000; 2000WO-US06456.
11-MAR-1999; 99US-0123888.
PR
```


Db 800 gaggggcagagcagagggggcaggggagggcagggagggagggagggagggg 859
Qy 430 gcgacagggcgccgagcaactgcccagcgctacgcccgcagcagcactgcccaagaaa 489
Db 860 caggaacagagggcgagggggcagggagcagggagggggcagggagggcgagcagag 919
Qy 490 ccggatgagaggaagagctcaagg 514
Db 920 gggcagagggggcagggagcagggag 944

RESULT 10

AAA75454

ID AAA75454 standard; DNA; 2580 BP.

XX AC AAA75454;

XX DT 15-JAN-2001 (first entry)

XX DE Nucleotide sequence of the Epstein Barr nuclear antigen.

XX KW Origin of replication; protein-protein interaction; replication;

XX KW two-hybrid system; nuclear antigen; ss.

XX OS Epstein-barr virus.

XX PN US6114111-A.

XX PD 05-SEP-2000.

XX PF 30-MAR-1998; 98US-0050863.

XX PR 30-MAR-1998; 98US-0050863.

XX PA (RIGE-) RIGEL PHARM INC.

XX PI Luo Y, Payan D, Huang B;

XX DR WPI; 2000-593546/56.

XX CC Composition for detecting protein-protein interactions in a mammalian
PT two hybrid system comprises bait and test vector which consist of
PT selection gene, vector viral origin of replication and fusion gene -

XX PS Disclosure; Column 17-20; 18pp; English.

XX CC The present sequence represents the Epstein Barr nuclear antigen.
CC It is used to produce bait vectors of the invention. The specification
CC describes a compositions and methods for a genetic system of detecting
CC protein-protein interactions in a mammalian host cell. The system
CC comprises bait and test, both containing selection genes, and viral
CC origin of replications which require bound viral replication proteins
CC to effect replication. The compositions is useful for detecting an
CC interaction between a bait protein and a test protein. It is useful in
CC a mammalian two-hybrid system for detecting protein-protein interactions
CC in a mammalian host cell.

XX SQ Sequence 2580 BP; 632 A; 512 C; 1054 G; 382 T; 0 other;

Query Match 2.9%; Score 78.6; DB 21; Length 2580;
Best Local Similarity 48.5%; Pred. No. 1.6e-08;
Matches 216; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 70 gaggatgccaccagaaagagggagggaggggtcgcgcgctctccgcccaag 129
Db 883 gaggggcagagcagagggggcagggagggagggggcagggggcagggagcag 942
Qy 130 cgcagcgcaaggcgcatcttcgggaagaagcccaagtcgcccccaagcagcccaag 189
Db 943 gaggagggcagagggggcagggagggagggagggggcagggagggggg 1002

Qy 190 ccggggaggaagaagggggtgccgagatgaagagagcccgctggaggacacgtgtgc 249
Db 1003 caggaggggcagggagggcgagggcagggagggcgagggagggcgaggggg 1062
Qy 250 gccgagagcccgagagagaggttgccatggcgagggcgagggagggcgagggagcc 309
Db 1063 caggagggggcagggagggcgagggagggcgagggagggcgagggagggcgaggg 1122
Qy 310 atgcagggagggaggttggttcggtcgcgcggggtcacccgggggaagaggggtgggaga 369
Db 1123 caggagggggcagggagggcgagggagggcgagggagggcgagggagggcgag 1182
Qy 370 aggaacgc 429
Db 1183 gaggggcagagcagggagggagggcgagggcgagggcgagggcgagggcgaggggg 1242
Qy 430 gcggacgagggcgcgagcaactggcccagcgctacggcgcgagcagcactgccgcaaa 489
Db 1243 caggagggggcagggagggcgagggagggcgagggagggcgagggagggcgaggg 1302
Qy 490 ccggatgagaggaagagctcaagg 514
Db 1303 gggcagagggggcagggagcagggag 1327

RESULT 11

AAX90923/C

ID AAX90923 standard; DNA; 5452 BP.

XX AC AAX90923;

XX XX 17-JAN-2000 (first entry)

XX DE Anti-sense strand of pCMVEBNA plasmid.

XX KW Anti-sense strand: plasmid pCMVEBNA; EBNA 1; episome; transfection;
KW Epstein Barr Virus Nuclear Antigen 1; origin of replication;
KW EBV oriP; eucaryotic host cell; recombinant cell line; ion channel;
KW multiple gene expression; receptor; transporter protein; gene therapy;
KW transcription factor; adhesion molecule; antisense therapy;
KW gene amplification; cell immortalisation; ss.

XX OS Epstein-barr virus.

XX OS Cytomegalovirus.

XX OS Synthetic.

XX FH Key Location/Qualifiers
FT CDS complement (3032..4957)
FT /*tag= a
FT /product= "EBNA 1 protein"
FT /note= "Epstein Barr Virus Nuclear Antigen 1"

XX WO947647-A1.

XX 23-SEP-1999.

XX XX 12-FEB-1999; 99WO-US03307.

XX XX 18-MAR-1998; 98US-0040961.

XX PR 06-AUG-1998; 98US-0130114.

XX XX (PHAR-) PHARMACOPETA INC.

XX PA Dama J BP, Horlick RA, Robbins AK;

XX XX WPI; 1999-010610/52.

XX DR New method for expressing genes from recombinant eukaryotic cells,
PT useful for gene therapy -
PT Example 1; Fig 1; 86pp; English.

XX PS The present sequence is an anti-sense strand of commercially available

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2002, 09:15:38 ; Search time 79.05 seconds
(without alignments)

Title: US-09-914-001-1

Perfect score:

Sequence: 1 atggcgccgagctcccgctc.....caggggaggtagttgagcag 2736

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 s

```
Database : Issued_Patents_NA:*
1: /cn2.6/ptdata/2/lna/5A_COMB.seq:*
2: /cn2.6/ptdata/2/lna/5B_COMB.seq:*
3: /cn2.6/ptdata/2/lna/6A_COMB.seq:*
4: /cn2.6/ptdata/2/lna/6B_COMB.seq:*
5: /cn2.6/ptdata/2/lna/PCTUS_COMB.seq:*
6: /cn2.6/ptdata/2/lna/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	93.2	3.4	1931	2	US-09-130-114-2	Sequence 2, Appli
C 2	79.4	2.9	7218	1	US-08-232-463-14	Sequence 14, Appl
C 3	78.6	2.9	2580	3	US-09-050-863-2	Sequence 2, Appli
C 4	78.6	2.9	2580	4	US-09-359-081-2	Sequence 2, Appli
C 5	78.6	2.9	5452	2	US-09-130-114-1	Sequence 1, Appli
C 6	78.6	2.9	9600	4	US-08-910-647-1	Sequence 1, Appli
C 7	78.6	2.9	10596	1	US-07-884-811-15	Sequence 15, Appl
C 8	78.6	2.9	10596	1	US-07-885-971-15	Sequence 15, Appl
C 9	78.6	2.9	10596	1	US-08-087-783A-15	Sequence 15, Appl
C 10	78.6	2.9	10596	1	US-08-194-088B-15	Sequence 15, Appl
C 11	78.6	2.9	10596	2	US-08-194-087-15	Sequence 15, Appl
C 12	78.6	2.9	10596	5	PCR-US93-04648-15	Sequence 15, Appl
C 13	76.2	2.8	51259	3	US-08-781-891-209	Sequence 209, App
C 14	75.4	2.8	16442	3	US-08-781-891-208	Sequence 208, App
C 15	70.4	2.6	3489	2	US-08-728-323A-1	Sequence 1, Appli
C 16	70.4	2.6	3207	2	US-08-770-379-20	Sequence 20, Appl
C 17	70.4	2.6	3207	4	US-08-757-669A-20	Sequence 20, Appl
C 18	70.4	2.6	3207	4	US-09-230-371A-20	Sequence 20, Appl
C 19	69.4	2.5	1548	2	US-08-762-106-5	Sequence 5, Appli
C 20	69.4	2.5	1548	4	US-09-320-774-5	Sequence 5, Appli
C 21	69.4	2.5	1581	2	US-08-762-106-6	Sequence 6, Appli
C 22	69.4	2.5	1581	4	US-09-320-774-6	Sequence 6, Appli
C 23	66.8	2.4	2277	1	US-08-676-967-5	Sequence 5, Appli
C 24	66.8	2.4	2277	1	US-08-676-974-5	Sequence 5, Appli
C 25	66.8	2.4	2277	2	US-09-098-487-5	Sequence 5, Appli
C 26	66	2.4	1771	2	US-08-533-669A-7	Sequence 7, Appli
C 27	66	2.4	1771	2	US-08-531-872-1	Sequence 1, Appli

ALIGNMENTS

RESULT

US-09-130-114-2/C

: Sequence 2, Application US/09130114

; Patent No. 5976807

; GENERAL INFORMATION:

APPLICANT: Horlick, Robert A.

APPLICANT: Damaj, Bassam B.

APPLICANT: Robbins, Alan K.

;
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes

; TITLE OF INVENTION: From Multiple Transfected Episomes

; FILE REFERENCE: 0867/1D903U

; CURRENT APPLICATION NUMBER: US/09/130,114

; CURRENT FILING DATE: 1998-08-06

; NUMBER OF SEQ ID NOS: 36

```
; SOFTWARE: FastSEQ for Windows Version 3.0
```

; SEQ ID NO 2

; LENGTH: 1931

; TYPE: DNA

; ORGANISM: EBNA

US-09-130-114-2

Query: Match 3.48: Score 93.2: PB 2: Length 1931:

Query: Match

Best Local Similarity 49.4%, P-Red: NO: 7.4e 14,
Matches 279: Conservative 0: Mismatches 283: Indels 6: Gaps

QY	44	ctggtgcggaagcgcgcgcgcgaagccgagagatccaccagaacaagaggagagagaggg	103
Db	984	CGGAGGTGGAGACGAGACGCGGGAGGACGAGGACGGGAGGAGGACGAGACGGGGAGG	925
QY	104	aggtctgcgcgcgcgtctctcccaagcgcgcgcgaagcgcatcttcctgcggaagaagc	163
Db	924	ACGGGAGGACGAGGACGCGGGAGGACGAGGACGGGAGGACGCGGGAGGACG	865
QY	164	ccaagtgcgccccccaagcagcgcccaagcggggagagaagaagaagggaatgccagatga	223
Db	864	AGGACGGGGAGACGGGGAGGAGGACGAGGACGGGGAGGACGGGGAGGAGGACGAGACG	805
QY	224	aggagccgctggaggacacgtgtgcgcgagagagccgcacagaggaggttggccatgg	283
Db	804	GGGAGGACGGGGAGGACGAGGACCGGGGAGGACGGGGAGGACGAGGACGGGGAGGACGGG	745
QY	284	gcgagagagagggccagagagcagcccatcagagagagtggtgtgcggtcgcgcgcgggt	343
Db	744	AGCAGCAGGACGGGGAGGACGAGGACCGGGAGGACGAGCAGCGGGGAGGACGAGGACGGGG	685
QY	344	caccgcggaagaagaggttggtgggaagaagcgcgcgcgcgcgcgcgcgcgcgcac----	399
Db	684	AGGACCGGGAGGACGGGGAGGAGGACGAGGACGAGGACGGGGAGGACGAGGACGGGGAGGACGGGG	625

Matches 216; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

```
QY 70 gagagatccaccagaacaagagagagaggtcgcgcgctctccgccaag 129
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 883 GAGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 942
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 130 cgcagcgcgaagcgcatcttcggggaagaagcccaagtcgcccccaagccaag 189
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 943 GAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1002
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 190 ccgggggaggaagaaggggatgcgagatgaagagcccggtggagacagcgtgtgc 249
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1003 CAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1062
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 250 gccagagagccagcagagagaggttgccatggcgagagagagcgagagcagcgc 309
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1063 CAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1122
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 310 atgcagagagaggtgttgcgtcgcgcggttcacccgggaagaagaggttgaggaga 369
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1123 CAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1182
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 370 aggaacgcgcgcgcgtggtgcacacagcgcgaggttcattcgagccctgtgtgc 429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1183 GAGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1242
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 430 gcgagcagcgcgagcaactgcccgaagctacggcgagcagcactgcgcaaaaga 489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1243 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1302
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 490 ccgagatgaggaagaagctcaagg 514
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1303 GGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1327
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 4

```
US-09-359-081-2
; Sequence 2, Application US/09359081
; Patent No. 6316223
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
;             Hiang, Betty
;             Pavan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
;                     System
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/359,081
; FILING DATE: 22-Jul-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/050,863
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-359-081-2
```

```
Query Match          2.9%; Score 78.6; DB 4; Length 2580;
Best Local Similarity 48.5%; Pred. No. 3.7e-10;
Matches 216; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 70 gagagatccaccagaacaagagagagaggttcgcgcgctctccgccaag 129
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 883 GAGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 942
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 130 cgcagcgcgaagcgcatcttcggggaagaagcccaagtcgcccccaagccaag 189
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 943 GAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1002
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 190 ccgggggaggaagaaggggatgcgagatgaagagcccggtggagacagcgtgtgc 249
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1003 CAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1062
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 250 gccagagagccagcagagagaggttgccatggcgagagagagcgagagcagcgc 309
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1063 CAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1122
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 310 atgcagagagaggtgttgcgtcgcgcggttcacccgggaagaagaggttgaggaga 369
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1123 CAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1182
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 370 aggaacgcgcgcgcgtggtgcacacagcgcgaggttcattcgagccctgtgtgc 429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1183 GAGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1242
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 430 gcgagcagcgcgagcaactgcccgaagctacggcgagcagcactgcgcaaaaga 489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1243 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1302
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 490 ccgagatgaggaagaagctcaagg 514
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1303 GGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1327
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
US-09-130-114-1/c
; Sequence 1, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; FROM MULTIPLE TRANSFECTED EPISOMES
; FILE REFERENCE: 0867/1D903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 5452
; TYPE: DNA
; ORGANISM: VEBNA
US-09-130-114-1
```

```
Query Match          2.9%; Score 78.6; DB 2; Length 5452;
Best Local Similarity 48.5%; Pred. No. 5.2e-10;
Matches 216; Conservative 0; Mismatches 229; Indels 0; Gaps 0;
```



```
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-884-811-15
```

```
Query Match          2.9%; Score 78.6; DB 1: Length 10596;
Best Local Similarity 48.5%; Pred. No. 7.1e-10;
Matches 216; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 70 gagagatccaccagaacaagaggagaggttcgcgcggtctcctccgccaag 129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2463 GAGGGGCAGGAGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGCAG 2522

QY 130 cgcagccgcaagcgcgcatcttcgggaagaagcccaagtcgcccccaagc 189
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2523 GAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGCAGGAG 2582

QY 190 ccggggaggaagaagaagggatcccgagatgaagagcccgctggagacagtgtgc 249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2583 CAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAG 2642

QY 250 gccgagagagcccgacgagaggttgccatggcgagagagagccgagagcgcc 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2643 CAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAG 2702

QY 310 atcgagagagaggttggttcggtcgcgcggttcacccgggaagaagaggggaga 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2703 CAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAG 2762

QY 370 aggaacgcgcgcgcgcgtggtgcgacacgagccgaggttcattcctggtgtgccc 429
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2763 GAGGGGCAGGAGCAGGAGGGGCAGGAGGGGCAGGAGGAGGGGCAGGAG 2822

QY 430 gcgagagagcgcgacgaactgccaagcgtacgcgcgcgcgcgcgcgcgcgcgc 489
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2823 CAGGAGCAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGGGCAGGAGGGGCAG 2882

QY 490 ccgagatgagaggaagagctcaagg 514
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2883 GGGCAGGAGGGGCAGGAGCAGGAGG 2907
```

```
RESULT 8
US-07-885-971-15
; Sequence 15, Application US/07885971
; Patent No. 5328637
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/885, 971
; FILING DATE: 19920518
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
```

```
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-885-971-15

Query Match          2.9%; Score 78.6; DB 1: Length 10596;
Best Local Similarity 48.5%; Pred. No. 7.1e-10;
Matches 216; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 70 gagagatccaccagaacaagaggagaggttcgcgcggtctcctccgccaag 129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2463 GAGGGGCAGGAGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGCAG 2522

QY 130 cgcagccgcaagcgcgcatcttcgggaagaagcccaagtcgcccccaagc 189
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2523 GAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGCAGGAG 2582

QY 190 ccggggaggaagaagaagggatcccgagatgaagagcccgctggagacagtgtgc 249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2583 CAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAG 2642

QY 250 gccgagagagcccgacgagaggttgccatggcgagagagagccgagagcgcc 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2643 CAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAG 2702

QY 310 atcgagagagaggttggttcggtcgcgcggttcacccgggaagaagaggggaga 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2703 CAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAG 2762

QY 370 aggaacgcgcgcgcgcgtggtgcgacacgagccgaggttcattcctggtgtgccc 429
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2763 GAGGGGCAGGAGCAGGAGGGGCAGGAGGGGCAGGAGGAGGGGCAGGAG 2822

QY 430 gcgagagagcgcgacgaactgccaagcgtacgcgcgcgcgcgcgcgcgcgcgc 489
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2823 CAGGAGCAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGGGCAGGAGGGGCAG 2882

QY 490 ccgagatgagaggaagagctcaagg 514
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2883 GGGCAGGAGGGGCAGGAGCAGGAGG 2907
```

```
RESULT 9
US-08-087-783A-15
; Sequence 15, Application US/08087783A
; Patent No. 5547856
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087.783A
; FILING DATE: 13-Jul-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/884811
; FILING DATE: 18-MAY-92
; PRIOR APPLICATION NUMBER: 07/885971
; FILING DATE: 18-MAY-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P0755779P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5416
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-087-783A-15
```

Query Match 2.9%; Score 78.6; DB 1; Length 10596;
Best Local Similarity 48.5%; Pred. No. 7.1e-10;
Matches 216; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

```
QY 70 gaggagatccaccagaacaggaggaggaggaggtcgcgcgccgttcctccgccaag 129
Db 2463 GAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2522
QY 130 cgcagccgaagcgcgcatcttcgggaagaagcccaagtcgcccccaagcagccaag 189
Db 2523 GAGGAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2702
QY 190 ccggggaggagaagaggggatgccagatgaaggagcccgctggaggagcagcgtgtgc 249
Db 2583 CAGGAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2642
QY 250 gccagagagcccgcccgctggtgcagccagcagcagcagcagcagcagcagcagcc 309
Db 2643 CAGGAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2702
QY 310 atgcaggaggaggttggttcggtcgcgcggttcacccggggaagagggtggggaga 369
Db 2703 CAGGAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2762
QY 370 aggaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 429
Db 2763 GAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2822
QY 430 gcgcagcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 489
Db 2823 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2882
QY 490 ccggatgaggaggaagagctcaagg 514
Db 2883 GGGCAGGAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2907
```

RESULT 10
US-08-194-088B-15
; Sequence 15, Application US/08194088B
; Patent No. 5580963
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
; TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS

```
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 Inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,088B
; FILING DATE: 09-FEB-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/884811
; FILING DATE: 18-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gallegos, R. Thomas
; REGISTRATION NUMBER: 32,692
; REFERENCE/DOCKET NUMBER: 755D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-2614
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-194-088B-15
```

Query Match 2.9%; Score 78.6; DB 1; Length 10596;
Best Local Similarity 48.5%; Pred. No. 7.1e-10;
Matches 216; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

```
QY 70 gaggagatccaccagaacaggaggaggaggagggtcgcgcgccgttcctccgccaag 129
Db 2463 GAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2522
QY 130 cgcagccgaagcgcgcatcttcgggaagaagcccaagtcgcccccaagcagccaag 189
Db 2523 GAGGAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2582
QY 190 ccggggaggagaagaggggatgccagatgaaggagcccgctggaggagcagcgtgtgc 249
Db 2583 CAGGAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2642
QY 250 gccagagagcccgcccgctggtgcagccagcagcagcagcagcagcagcagcc 309
Db 2643 CAGGAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2702
QY 310 atgcaggaggaggttggttcggtcgcgcggttcacccggggaagagggtggggaga 369
Db 2703 CAGGAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2762
QY 370 aggaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 429
Db 2763 GAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2822
QY 430 gcgcagcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 489
Db 2823 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2882
QY 490 ccggatgaggaggaagagctcaagg 514
Db 2883 GGGCAGGAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2907
```


